Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Signal Processing in Acoustics
Session 4pSP: Sampling Methods for Bayesian Analysis and Inversions in Acoustic Applications

4pSP4. Using nested sampling with Galilean Monte Carlo for model comparison problems in acoustics

Paul Goggans*, R. W. Henderson and Ning Xiang

***Corresponding author's address: Electrical Engineering, University of Mississippi, Anderson Hall Room 302, University, MS 38677, goggans@olemiss.edu**

Nested sampling is increasingly being used to calculate the evidence for competing models in Bayesian model comparison problems arising in acoustics applications. Use of nested sampling offers advantages in robustness over alternative methods of evidence calculation and enables evidence calculation for models with many parameters. The most challenging aspect of implementing nested sampling is sampling from the prior for the parameters constrained by a threshold likelihood value. For models with just a few parameters, sampling from the constrained prior can be accomplished with a simple Monte Carlo algorithm implementing a random walk, however, this simple method is inefficient and fails as the number of model parameters increases. John Skilling, the originator of nested sampling, has proposed the "Galilean" Monte Carlo method for efficiently sampling from the constrained prior when there are many parameters. Unlike the random walk method, the Galilean Monte Carlo method moves samples with a vector velocity, reflecting them from the likelihood constraint surface when necessary. This directed sampling gives the method its greater efficiency. In this paper we discuss our experience in implementing Galilean Monte Carlo in nested sampling and compare Galilean and random walk Monte Carlo for a model comparison problem in room acoustics.

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4800876]
Received 22 Jan 2013; published 2 Jun 2013
Proceedings of Meetings on Acoustics, Vol. 19, 055089 (2013)

Nested Sampling in Acoustics

In the domain of acoustics, many problems can be solved effectively through the judicious use of Bayesian inference. Examples of such problems include multiple decay slope analysis [1, 2], room acoustics modal analysis [3], and estimation of acoustic parameters in porous materials [4]. This paper will take modal analysis as an example to discuss the use of nested sampling and related techniques in acoustic applications.

The acoustical signature of a room is represented by the room's acoustic impulse response. This impulse response is dependent on the room's modal characteristics, determined by the room's geometry and the reflective characteristics of the room's surfaces. The geometry of the room determines the frequencies of the room's resonant modes, while the surface characteristics determine the decay times and amplitudes of these modes.

The room impulse response can be modeled as a sum of decaying sinusoids:

$$\gamma(t) = \sum_{i=1}^{K} \exp(-6.9t/\tau_i)[a_i \cos(2\pi f_i t) + b_i \sin(2\pi f_i t)]. \tag{1}$$

Each individual decaying sinusoid corresponds to a resonant mode in the room.

Two levels of Bayesian inference are used to perform room mode analysis: model comparison and parameter estimation. Model comparison allows the number of modes present to be inferred. Measured data are modeled with different numbers of modes, and these competing models are ranked using Bayesian evidence. Once a model is selected, parameter estimation is performed to estimate the values of the model parameters (amplitudes, decay times, and frequencies).

Various methods exist for performing both levels of inference; however, the potentially large number of parameters present in this problem indicates that a method that elegantly deals with high dimensionality would be ideal. Nested sampling [5, 6, 7] is such a method. Nested sampling not only deals robustly with a large number of parameters, but it also provides a straightforward method for calculating Bayesian evidence needed for model comparison. Nested sampling has been used in this context in previous work [3].

Nested sampling (Algorithm 1) begins by drawing a population of samples from the unconstrained prior distribution. In this case, the model is reparameterized such that its prior distribution is uniform on the unit hypercube. The likelihood values of these samples are evaluated, and the sample with the lowest likelihood is discarded into a separate collection of samples. The likelihood of this discarded sample is used as a likelihood constraint. The discarded sample is replaced by sampling from the prior distribution constrained by the likelihood constraint. Any method that generates independent samples from the constrained prior distribution may be used to replace the discarded sample; however, a Markov-Chain Monte Carlo (MCMC) process is generally used for this purpose.

Algorithm 1 Nested Sampling

Draw N samples $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_N$ from the uniform prior distribution
Calculate likelihood $\mathscr{L}_i$ for each sample.
$Z \leftarrow 0$ ▷ Initialize evidence at 0
$X_0 \leftarrow 1$ ▷ Initialize prior mass at 1
for $i \leftarrow 1, M$ **do**
  $\mathscr{L}^* \leftarrow \texttt{min}(\mathscr{L})$ ▷ Minimum likelihood becomes likelihood constraint
  $X_i \leftarrow \exp(-i/N)$
  $w_i \leftarrow X_{i-1} - X_i$ ▷ Prior mass width
  $Z \leftarrow Z + \mathscr{L}^* w_i$
  $\phi_i \leftarrow \theta_*$ ▷ Discard lowest-likelihood sample and save in separate array
  Choose j from uniform distribution of integers on $[1, *) \cup (*, N]$ ▷ Pick a random index
  $\theta_* \leftarrow \texttt{MCMC}(\theta_j, \mathscr{L}^*)$ ▷ Replace discarded sample using MCMC
end for

As the nested sampling process iterates, the prior mass at each step is calculated. The prior mass is equal to the integral of the constrained prior distribution over the parameters. Also at each iteration, the

contribution of the sample to the evidence is calculated as the product of the change in prior mass and the likelihood constraint. The nested sampling process continues until the desired degree of convergence is reached. After completion of the nested sampling algorithm, the values of the model parameters can be estimated by finding the weighted mean of the discarded samples with respect to the parameters.

A significant challenge to effectively implementing nested sampling lies in choosing a robust method for drawing independent samples from the constrained prior distribution. The remainder of this paper will describe two choices for fulfilling this role: random walk MCMC [8] and a newer Monte Carlo method, Galilean Monte Carlo [9].

Random-walk Markov Chain Monte Carlo

A random-walk implementation [10] (Algorithm 2) of Metropolis Hastings MCMC is used in [3] to generate new prior samples in the nested sampling process. The MCMC process uses a randomly chosen surviving sample from the working population of samples to begin the chain. For the first step, a single parameter is chosen at random to vary. This parameter is increased or decreased by a value drawn from a uniform distribution on $[-0.1, 0.1]$. If the likelihood of the sample with the varied parameter value is greater than the likelihood constraint, the new sample is accepted. If the likelihood of the sample is less than the likelihood constraint, the new sample is rejected, and the previous chain step is copied into the current chain position. At each iteration, the total numbers of proposed moves accepted and rejected are compared. If there are more moves accepted than rejected, the step size is decreased. If more moves are rejected than accepted, the step size is increased. A parameter is chosen at random once again, and the process is repeated until a predetermined number of elements have been added to the chain. The final chain element is returned as the new sample.

Random walk is straightforward and a computationally inexpensive way to generate a new sample. However, as the size of the parameter space increases, it has an increasingly difficult time fully exploring the the prior distribution constrained by the likelihood. In this case the final chain step may not be independent of the initial point so that the the method fails to provide an independent sample from the constrained prior as required by nested sampling. This disadvantage becomes especially apparent as the number of parameters becomes large (such as in this room mode analysis problem when there are many modes) and results in poor estimates for the evidence and the parameter values. This leads us to search for a more robust method.

Galilean Monte Carlo

Galielan Monte Carlo [9] is an MCMC method developed by John Skilling, the author of the original nested sampling algorithm [5]. This method addresses the disadvantages of random walk MCMC listed previously: the constrained prior can be explored thoroughly, and the method is well-suited to many-dimensional distributions.

Similarly to random walk, the Galilean Monte Carlo algorithm (Algorithm 3) begins with a randomly selected surviving sample from the working population. This sample is treated as a particle, with the sample's parameter values denoting the particle's position, $\mathbf{x}$. The particle's velocity is initialized as

$$\mathbf{v} = \mathbf{S}\mathbf{r}, \tag{2}$$

where $\mathbf{r}$ is a vector with values drawn from $\texttt{Normal}(\mathbf{0}, \mathbf{I})$ and $\mathbf{S}$ is a semimetric matrix that conditions the velocity in a useful fashion.

A particularly useful way to define $\mathbf{S}$ is as follows:

$$\mathbf{S}\mathbf{S}^t \approx (-\nabla\nabla \log L)^{-1}, \tag{3}$$

where $\nabla\nabla \log L$ is the Hessian (curvature) with respect to the parameters of the log-likelihood function, and, if the likelihood function is a multivariate Gaussian, is also the negative inverse of the covariance matrix. Using this matrix to initialize the velocity causes Galilean MC to view the constrained space as a

Algorithm 2 Random Walk

function RANDOMWALKMC($\mathbf{x}, \mathscr{L}^*$)
  $N \leftarrow 41$ ▷ Number of chain steps
  $L \leftarrow$ `length`($\mathbf{x}$)
  $a \leftarrow 0$ ▷ Number of acceptances
  $r \leftarrow 0$ ▷ Number of rejections
  $\mathbf{mc} \leftarrow \mathbf{x}$ ▷ First element of chain is initial sample
  for $i \leftarrow 2, N$ **do**
    $j \sim$ `Uniform`$(1, L)$ ▷ Integers only. Pick random index of sample vector to vary
    $z \sim$ `Uniform`$(-0.1, 0.1)$ ▷ Step size
    $\mathbf{y} \leftarrow \mathbf{mc}$ ▷ Initialize current chain step with previous step
    $\mathbf{y}_j \leftarrow z + \mathbf{mc}_j$
    $\mathscr{L} \leftarrow$ `logL`($\mathbf{y}$)
    if $\mathscr{L} > \mathscr{L}^*$ **then**
      $a \leftarrow a + 1$ ▷ Increment number of acceptances
      $\mathbf{mc} \leftarrow \mathbf{y}$
    else
      $r \leftarrow r + 1$ ▷ Increment number of rejections
    end if
    if $a > r$ **then**
      $z \leftarrow z \times \exp(1/a)$ ▷ If acceptance ratio > 50%, shrink step size
    else if $a < r$ **then**
      $z \leftarrow z / \exp(1/r)$ ▷ If acceptance ratio < 50%, increase step size
    end if
  end for
  return mc
end function

Algorithm 3 Galilean Monte Carlo

function GMC($\mathbf{x}, \mathscr{L}^*$)
  $\tau \leftarrow 1$
  $\mathbf{S} \leftarrow (-\nabla\nabla \mathtt{logL}(\mathbf{x} + [\text{some perturbation}]))^{-1/2}$
  $\mathbf{r} \sim \mathtt{Normal}(\mathbf{0}, \mathbf{I})$ ▷ $\mathbf{r}$ is same length as $\mathbf{x}$
  $\mathbf{v} \leftarrow \mathbf{S}\mathbf{r}$ ▷ Initial velocity
  Adjust τ such that the first movement of $\mathbf{x}$ is inside boundaries.
  for $i \leftarrow 1, 100$ **do**
    $\mathbf{x}' \leftarrow \mathbf{x} + \tau\mathbf{v}$ ▷ Calculate position after proceeding
    $\mathscr{L}' \leftarrow \mathtt{logL}(\mathbf{x}')$ ▷ Determine log-likelihood at new position
    if $\mathbf{x}'$ is outside prior boundaries OR $\mathscr{L}' \le \mathscr{L}^*$ **then**
      if $\mathbf{x}'$ is outside prior boundaries **then**
        $\mathbf{g} \leftarrow$ Unit vector normal to boundary
      else if $\mathscr{L}' \le \mathscr{L}^*$ **then**
        $\mathbf{g} \leftarrow \nabla \mathtt{logL}(\mathbf{x}')$
      end if
      $\mathbf{v}' \leftarrow \mathbf{v} - 2\mathbf{S}\mathbf{S}^t\mathbf{g}\frac{\mathbf{g}^t\mathbf{v}}{\mathbf{g}^t\mathbf{S}\mathbf{S}^t\mathbf{g}}$ ▷ Calculate specular reflection velocity
      $\mathbf{x}'' \leftarrow \mathbf{x} + \tau\mathbf{v} + \tau\mathbf{v}'$
      $\mathscr{L}'' \leftarrow \mathtt{logL}(\mathbf{x}'')$
      if $\mathbf{x}''$ is outside prior boundaries or $\mathscr{L}'' \le \mathscr{L}^*$ **then**
        $\mathbf{x}' \leftarrow \mathbf{x} - \tau\mathbf{v}$ ▷ Try reversing the original direction
        $\mathbf{x} \leftarrow \mathbf{x}'$ ▷ Set as new position
        $\mathbf{v} \leftarrow -\mathbf{v}$ ▷ Set as new velocity
      else ▷ If $\mathbf{x}''$ is OK...
        $\mathbf{x} \leftarrow \mathbf{x}''$
        $\mathbf{v} \leftarrow \mathbf{v}'$
      end if
    else
      $\mathbf{x} \leftarrow \mathbf{x}'$ ▷ Proceed to new position
      $\mathbf{accept}_i \leftarrow 1$ ▷ Acknowledge that new point was accepted
    end if
    if $i > 20$ **then**
      $accept\_ratio \leftarrow \sum_{j=i-20}^{i} accept_i/20$
    end if
    if $accept\_ratio \le 0$ **then**
      $accept\_ratio \leftarrow 1/20$
    end if
    $\tau \leftarrow \tau\left(\frac{accept\_ratio}{0.9}\right)^{(1/20)}$
  end for
end function

hypersphere, helping to achieve even coverage of the space, even if the space has awkwardly-shaped boundaries. This definition yields

$$\mathbf{S} \approx (-\nabla\nabla \log L)^{-1/2}. \tag{4}$$

In this particular problem, an analytic evaluation of the Hessian is impractical, and finite-difference approximations are both computationally expensive and inadequate. However, clever use of conjugate gradient methods as described in [11] can produce a useful value for $\mathbf{S}$ with minimal computational cost. In order to define $\mathbf{S}$ in such a way as to maintain detailed balance throughout the motion of the particle, it should not be calculated at the initial position of the particle; however, to be most useful, it should be calculated at a point within the same likelihood mode as the particle. Such a position can be determined by moving the particle one step with a random velocity not equal to $\mathbf{v}$ and ensuring that the resultant point is within the likelihood and prior boundaries.

With the initial position and velocity determined, the particle can begin its motion. The basic idea is that the particle moves about the prior space constrained by the likelihood and specularly reflects off of the likelihood or prior boundaries whenever it strikes them. The equation describing the particle's initial step is

$$\mathbf{x}' = \mathbf{x} + \tau\mathbf{v}, \tag{5}$$

where τ is the time step and $\mathbf{x}'$ is the particle's new position, assuming it does not strike a boundary. The time step τ is set at the beginning of the process so that a step of the particle is on the same order of magnitude as the size of the space inside the boundaries. τ may be adjusted throughout the process to maintain a certain acceptance ratio. When a proposed particle position $\mathbf{x}'$ lies outside of the likelihood constraint, the gradient of the likelihood is calculated at that point,

$$\mathbf{g} = \nabla \mathtt{logL}(\mathbf{x}'), \tag{6}$$

indicating the correct adjustment for the velocity so that the chain particle reflects back inside the volume where the likelihood constraint is satisfied. If the proposed particle instead lies outside of the prior distribution's boundaries, the gradient is simply defined as the vector normal to the prior boundary. In both cases, the reflection follows the equation

$$\mathbf{x}'' = \mathbf{x} + \tau\mathbf{v} + \tau\mathbf{v}' \tag{7}$$

with

$$\mathbf{v}' = \mathbf{v} - 2\mathbf{S}\mathbf{S}^t\mathbf{g}\frac{\mathbf{g}^t\mathbf{v}}{\mathbf{g}^t\mathbf{S}\mathbf{S}^t\mathbf{g}} \tag{8}$$

and reflected position $\mathbf{x}''$.

Occasionally, the shape of the likelihood boundary can cause the reflected position, $\mathbf{x}''$, to lie outside of the likelihood boundaries. In this case, the motion of the particle is reversed from its initial position, yielding

$$\mathbf{x}' = \mathbf{x} - \tau\mathbf{v}. \tag{9}$$

Once a particle position that satisfies all prior and likelihood requirements is found, it is accepted as the new position. The particle moves a predetermined number of times, and the final position is used as the new sample for nested sampling.

CONCLUSION

Room mode analysis presents an interesting challenge to Bayesian inference with nested sampling. The large number of parameters involved in the analysis of rooms with many modes precludes use of more traditional random-walk Monte Carlo and motivates use of a more robust method to explore the prior distribution. Galilean Monte Carlo presents an alternative that shows promise in dealing with the high dimensionality of room mode analysis. Preliminary work indicates that Galilean Monte Carlo can produce results superior to random walk in room mode analyses involving many modes. More work is needed to refine our implementation of the algorithm, and our final conclusions will be presented at the meeting in Montreal.

References

[1] N. Xiang and P. Goggans, "Evaluation of decay times in coupled spaces: Bayesian parameter estimation", J. Acoust. Soc. Am. **110**, 1415 (2001).

[2] N. Xiang and P. Goggans, "Evaluation of decay times in coupled spaces: Bayesian decay model selection", J. Acoust. Soc. Am. **113**, 2685 (2003).

[3] W. Henderson, P. Goggans, N. Xiang, and J. Botts, "Bayesian inference approach to room-acoustics modal analysis", Bayesian Inference and Maximum Entropy Methods in Science and Engineering (In Print) (2013).

[4] C. Fackler, E. Dieckman, and N. Xiang, "Porous material parameter estimation: A Bayesian approach", AIP Conference Proceedings **1443**, 314–321 (2012).

[5] J. Skilling, "Nested sampling", AIP Conference Proceedings **735**, 395–405 (2004).

[6] J. Skilling, "Nested sampling for general Bayesian computation", Bayesian Analysis **1**, 833–860 (2006).

[7] M. Betancourt, "Nested sampling with constrained Hamiltonian Monte Carlo", AIP Conference Proceedings **1305**, 165–172 (2011).

[8] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, eds., *Markov Chain Monte Carlo in Practice* (Chapman & Hall/CRC) (1996).

[9] J. Skilling, "Bayesian computation in big spaces: Nested sampling and Galilean Monte Carlo", AIP Conference Proceedings **1443**, 145 (2012).

[10] D. S. Sivia and J. Skilling, *Data analysis: a Bayesian tutorial* (Oxford University Press, USA) (2006).

[11] J. Skilling, "Conjugate gradient for Bayesian computation", AIP Conference Proceedings **1193**, 269–276 (2009).