Project Paper Update

Tracey Marsh

Group Health Research Institute
University of Washington, Department of Biostatistics

Paper

Estimating and Projecting Trends in HIV/AIDS Generalized Epidemics Using Incremental Mixture Importance Sampling

Adrain E. Raftery and Le Bao

Biometrics 2010

Method - IMIS

Incremental Mixture Importance Sampling

- numerical algorithm for sampling from a posterior
- addresses limitations of current algorithm
  Sampling Importance Resampling (SIR)
  posteriors with multimodality & nonlinear ridges

Bayesian framework in the context of modeling

- Model: $\theta \xrightarrow{M} \rho$
  - M deterministic scientific model
    population dynamics model for bowhead whales
  - θ input parameters for model (settings)
    birth and death rates, initial population size, ...
  - $\rho = M(\theta)$ model output
    population size by year
- Model Calibration (select θ)
  - prior on input parameters: $\pi(\theta)$
  - X is observed real data
    for certain years: observed population counts
  - posterior on input parameters: $\pi(\theta|X)$
    use $\mathbb{E}[\theta|X]$ for point estimates of ρ
    samples from $\pi(\theta|X)$ for intervals around $\widehat{\rho}$
- Bayes Formula: $\pi(\theta|X) = \frac{L(X|M(\theta))\pi(\theta)}{L(X)}$

Sampling Importance Re-sampling

- sample from prior: $\{\theta_1, \ldots, \theta_N\} \sim \pi_0$
- run model to get output for each: $\{M(\theta_1), \ldots, M(\theta_N)\}$
- calculate likelihood of model output: $L_i \equiv L(X|M(\theta_i))$
- calculate importance weight: $\omega_i = \frac{L_i}{\sum_j L_j}$
- weighted re-sample of $\theta_1, \ldots, \theta_N$

  estimate of posterior: $\pi(\theta|X)$

Note: unique points of posterior will always be a subset of the unique points sampled from prior.

Incremental Mixture Importance Sampling

- Initial sample and weights: $\left\{\left(\theta_i, \omega_i = \frac{L_i}{\sum_j L_j}\right) 1 \leq i \leq N_0\right\}$
  resampling now gives SIR estimate of posterior
- 'fill-in' important regions: $1 \leq k \leq K$
  identify underrepresented neighborhood
  add B points of Normal mass
  use mixture distribution as new prior: π_k
  update weights: ω_i^k
- repeat until stopping criteria met: K times
  expected % unique points in resample $\geq 1 - 1/e$
- weighted re-sample from $\left\{\left(\theta_i, \omega_i^K\right) 1 \leq i \leq N_0 + KB\right\}$
  estimate of posterior: $\pi(\theta|X)$

IMIS - Add Normal Mass

Expand sample with B points sampled from $q_k = N\left(\theta^k, \Sigma^k\right)$

- $\theta^k = \arg\max_{\theta_i}\{\omega^{k-1}(\theta_i)\}$
  center of important neighborhood
- Σ^k
  weighted covariance of B points of current sample
  in the neighborhood of θ^k
  Mahalanobis metric w.r.t. π_0
  weights $\propto \omega_i + 1/N_k$

$S_k = S_{k-1} \cup \{\theta_{k,1} \ldots \theta_{k,B}\}$
$N_k \equiv \#S_k = N_0 + kB$

Side Note - Mahalanobis Metric

$$d(\theta_i, \theta^k) = \sqrt{(\theta_i - \theta^k)^T \Sigma^{-1} (\theta_i - \theta^k)}$$

Σ is covariance matrix for π_0

Can be thought of as a dissimilarity measure between two points of the same distribution with covariance Σ.

Using this to grab all the B points of the sample that form the narrowest percentile range centered at θ^k.

IMIS -Update weights

At end of iteration k,

mixing sampling distribution:

- $\pi_k(\theta) = N_k^{-1}\left(N_0\pi_0(\theta) + B\sum_j^k q_k(\theta)\right)$

weights:

- $\omega_i^k \propto L_i(X|M(\theta_i)) \times \underbrace{\pi_0(\theta_i)\pi_k(\theta_i)^{-1}}_{adj}$

  away from important spots: $adj \approx L_i$
  near θ^k: $adj \ll 1$
  $\pi_k = \pi(\theta|X)$, then $\omega_i \propto 1$

Replication Goal

Reproduce results for two methods of interest from simulation study.

Scenario	SIR	IMIS
Ridge-Like	2e-5	0.0675
Bimodal	0.0002	0.2063

Table : ESS$/N_K$

- Prior, model, and likelihood of model output specified
- Evaluated in terms of efficiency: $\frac{ESS}{\#evaluations}$

Simulation - Set-up

Scenario: Ridge-Like

Model:

$$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6) \xrightarrow{M} \left(\prod_{i=1}^{4} \theta_i, \theta_2\theta_4, \frac{\theta_1}{\theta_5}, \theta_3\theta_6 \right)$$

Prior:

$$\pi_0(\theta) \sim N(\mu_0, \mathrm{Diag}(\sigma_0)) \qquad \mu_0, \sigma_0 \text{ specified}$$

Likelihood:

$$L(X|M(\theta)) \sim N(\mu_L, \mathrm{Diag}(\sigma_L)) \qquad \mu_L, \sigma_L \text{ specified}$$

Methods - Evaluation

Evaluated in terms of efficiency: $\frac{ESS}{N_K} \leq 1$

- ESS $=$ effective sample size of N_K indirect samples
- compared to N_K direct samples from $\pi(\theta|X)$

Effective sample size [Kong JASA 1994]:

$$
\begin{aligned}
\text{ESS} &= \frac{N_K}{1 + CV} \\
&= \frac{N_K}{1 + \frac{VAR[\omega]}{\mathbb{E}^2[\omega]}} \\
&= \frac{N_K}{\frac{\mathbb{E}[\omega^2]}{\mathbb{E}^2[\omega]}} \\
\widehat{\text{ESS}} &= \frac{1}{\sum_i \omega_i^2} \qquad \widehat{\mathbb{E}[\omega]} = 1/N_k
\end{aligned}
$$

Next Steps

- stopping criteria - understand

  expected % of unique points $\geq 1 - 1/e$

  value when weights all equal
- simulation

  numerical under/over-flow issues

  covariance estimates - not positive definite

  empirically approximate stopping criteria