STATS 700-002 Class 10.
Variational phylodynamic inference

Aaron King and Edward Ionides

October 30, 2025

Outline

1. How is variational inference applied to phylodynamic models?

2. The role of automatic differentation.

Bayesian vs non-Bayesian phylodynamic inference

▶ To what extent do you agree with the assertion that phylogenetic uncertainty is not reflected in frequentist point estimates but is included in Bayesian estimates?

Variational inference

Fourment M, Darling AE. 2019. Evaluating probabilistic programming and fast variational Bayesian inference in phylogenetics. PeerJ. 7:e8272.

Fourment M, Magee AF, Whidden C, Bilge A, Matsen FA IV, Minin VN. 2020. 19 dubious ways to compute the marginal likelihood of a phylogenetic tree topology. Syst Biol. 69(2):209–220.

Zhang C. 2020. Improved variational Bayesian phylogenetic inference with normalizing flows. Advances in neural information processing systems, Vol. 33. p. 18760–18771.

Zhang C, Matsen FA IV. 2019. Variational Bayesian phylogenetic inference. In 7th International Conference on Learning Representations, ICLR 2019; 2019 May 6–9

Approximations

▶ "A divide-and-conquer strategy where distant subtrees of a very large phylogeny are assumed to evolve approximately independently."

▶ "Assume that topological estimates of these subtrees are an accurate reflection of their distribution under the prior."
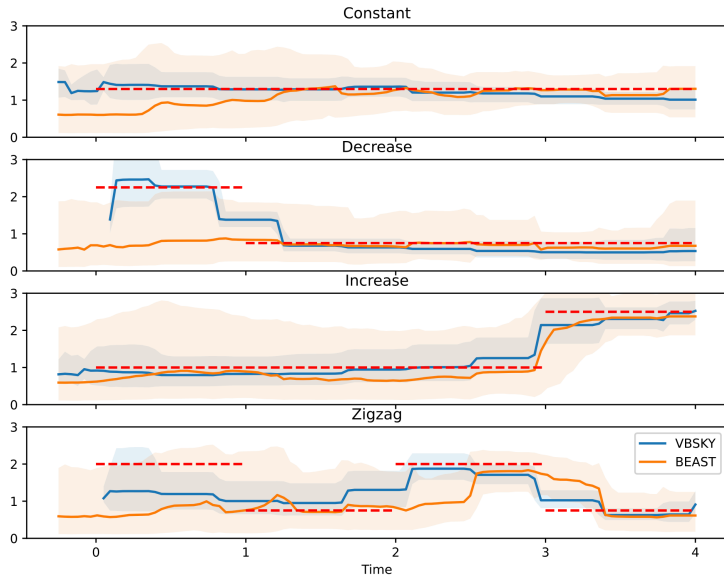
Fɪɢ. 1. Median of the medians and the equal-tailed 95% credible intervals of the posteriors of the effective reproductive number over time of the 10 simulations for each scenario using VBSKY and BEAST. The dotted line is the true effective reproductive number over time.

Fig 1

▶ This is to validate VBSKY on a problem that BEAST can also handle. VSBKY can handle many problems outside the scale of BEAST, and in that case we cannot make direct comparison.

▶ Which method is more accurate in these examples? How do you determine this?

Fig. 2. Posterior of *R* for Florida, Michigan, and the USA using biased sampling and a strong prior on *s*. For each method the posterior median and equal-tailed 95% credible interval are shown. The dotted line is $R = 1$.

Fig 2

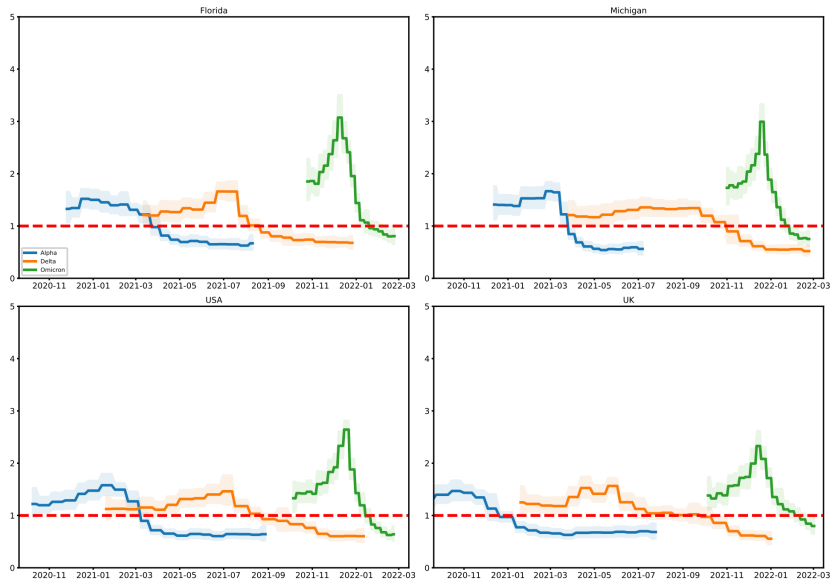▶ What is *biased sampling* and how/why was it done?

FIG. 3. The posterior median and equal-tailed 95% credible interval of *R* for the Alpha, Delta, Omicron variants.

Fig 3, Strain dynamics

The model has a prior on R that is time-invariant. For example, "Since $R > 1$ is not a reasonable prior assumption for a strain which is known to have vanished, we shrank the prior distribution towards zero to attenuate this effect." What is the scientific challenge for interpreting inferred R values to learn about the biology?

Molecular clocks

The paper uses a "strict molecular clock" model (page 9). What is this, and what are the consequences of that decision?

Heuristics

▶ Developing demonstrably good approximations that enable quantitative analysis of large populations is a useful scientific task. Statistical mechanics for ideal gasses is a classic example.

Heuristic 1. Conditional independence of disjoint parts of the tree. Here, we take random sub-samples.

Heuristic 2. Topology is approximately independent of the sequence data, given the parameters. We can separate estimation of the topology (from any scalable tree estimation method) from estimation of the branch lengths.

▶ Separating the topology provides a continuous problem, amenable to a variational approximation.

Variational inference

▶ Key idea: finding the closest approximation to a Bayesian posterior within a tractable model class can be much easier than finding the exact posterior.

▶ This approximation problem involves maximization, not Monte Carlo sampling.