

Discussion: Convergence Rate of Markov Chain Methods for Genomic Motif Discovery

Bala Rajaratnam
Stanford University

ADAPSKIII, January 2011

(The Canyons, Park City, Utah)

General Remarks

- Motif discovery is a fundamental problem in genomics and biomedical research. Extensive literature.
- As pointed out by the speaker, motif discovery is also a difficult problem.
- The work undertaken by the authors is therefore very important.
- Different MCMC algorithms (and modifications) have been proposed in the literature for motif discovery.
- The work undertaken by the authors establishes theoretical properties of a popular Gibbs sampling method.
- Theoretical safeguards are equally critical in high dimensional statistical applications.

General Remarks

- The research problem is important as it studies a widely used, applicable and representative Gibbs sampler.
- Despite heroic efforts in the literature, this work is one of the very few that consider bounds on convergence rates with a real statistical application in mind. Challenging problem !
- The work has a theoretical message with implications for applications and practical settings.
- From a theoretical perspective geometric ergodicity, or even better, uniform ergodicity can still lead to poorly behaved samplers.
- The results reiterates the fact that when using methods proposed in the literature, one should proceed with caution.

General Remarks

- Analysis of bounds on convergence rates for MC on discrete state spaces
- Spectral Analysis approach
- Behavior of the second largest eigenvalue of transition kernel as a function of the dimension L .
- When there are multiple true motifs, the theory presented by the authors paints a bleak picture.
- Scalability of the Gibbs algorithm is called into question.
- Nevertheless (as pointed out by the authors) it is useful for (posterior) mode-finding.

Questions for Presenter

1 **S**: sequence of nucleotides $\mathbf{S} \in \{1, 2\}^L$ used for illustration. If $\mathbf{S} \in \{1, 2, \dots, r\}^L$ what role (if any) would r have on mixing results.

2 Priors:

- $\theta_j \sim \text{Dirichlet}(\beta_j); j = 0, \dots, w$
- $\mathbf{A}_j \stackrel{iid}{\sim} \text{Bernoulli}(p_0)$ for fixed p_0 .

How does one choose hyperparameters β_j ?

3 Does fixing p_0 not strongly dictate or have an impact on the proportion of sites where motifs end ? How does the data generating mechanism overcome this ?

4 Would it make sense to to use a hierarchical approach, e.g. use a uniform prior for p_0 ? Presumably the added layer presents more difficulty in the analysis ?

5 The Gibbs sampler is problematic in terms of mixing time. It should still be better than frequentist approaches, e.g., the EM algorithm ? since Gibbs more flexible than the "greedy" EM ? Any comments/insights ?

6 "Sampler gets stuck in local modes". Scope for adaptive MCMC ?

Concluding Remarks

- Authors are to be commended without reservation for an interesting piece of work.
- Paper presents theory that is well motivated by applications.
- The work has potential implications for other Gibbs samplers in use in high dimensional statistical inference.