Gibbs PRAMpling

Parallel Gibbs Sampling Methods for Gaussian Mixture Models

Henry Ehrenberg

Gibbs sampling: MCMC method to approximate difficult distributions

Idea: Iteratively update each variable conditioned on the others

---

Algorithm 1: Gibbs sampling (linear scan, sequential)

Data: Variables $\{z_m\}_{m=1}^{M}$, distribution D, iterations T

begin

  for $t = 1$ *to* T **do**

    for $m = 1$ *to* M **do**

      Re-sample z_m from $P_D(Z_m | Z_{-m})$

Parallelize through MC properties, connections to SGD

Gibbs sampling for Gaussian mixture models (GMMs)

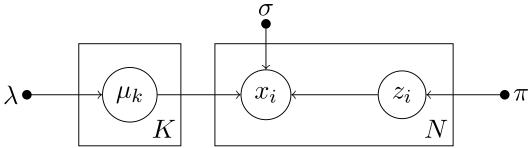

Figure: Generative model for GMMs

$$\mu_k \sim \mathcal{N}(0, \lambda^2 I)$$
$$z_i \sim \mathbf{Cat}(\pi)$$
$$x_i \sim \mathcal{N}(\mu_{z_i}, \sigma^2 I)$$

Algorithm 2: Gibbs sampling for GMMs (linear scan, sequential)

Data: Data $\{x_i\}_{i=1}^N$, parameters π, σ^2, λ^2

begin

    Initialize $\hat{z}, \hat{\mu}$
    for $t = 1$ *to* T **do**
        for $i = 1$ *to* N **do**
            Re-sample $\hat{z}_i$ from $P(z_i|\pi, \hat{\mu}, x_i)$
        for $k = 1$ *to* K **do**
            Re-sample $\hat{\mu}_k$ from $P(\mu_k|\hat{z}, x)$

Analysis of PRAM Gibbs sampling methods

Chromatic GS: Sample independent sub-chains in parallel

HogWild! GS: Sample all variables asynchronously

Iteratively warmer starts GS: Aggregate independent samplers

Algorithm	Work	Depth	Improvement
Sequential GS	$O(TNKd)$	$O(TNKd)$	None
Chromatic GS	$O(TNKd)$	$O(T(Kd + Nd))$	μ_k, z_i parallel
HogWild! GS	$O(TNKd)$	$O(Kd + Nd)$	All parallel
IWS GS	$O(TNKd)$	$O(TNKd)$	Initialization

Thanks to Alex Ratner, Bryan He, and Chris Ré

References

De Sa et al. *Ensuring Rapid Mixing and Low Bias for Asynchronous Gibbs Sampling*. ICML '16.

Blei. *Bayesian Mixture Models and the Gibbs Sampler*. Foundations of Graphical Models '15.