Abstract:

1. Abstract implies that items (1) and (2) are different but in fact non-linear constraints are an example of the case where piecewise distributions emerge.

2. It seems that Gibbs in particular has problems with (1) and (2) that we address here but in fact all MCMCs have problems not just Gibbs.

In intro:

1- Same problems as in abstract.

2- HMC (and its first ref) should be deleted to prevent confusion.

3- It is quite confusing where we speak about piecewise distributions and where about the deterministic relations. We go back and forth.

4- Why we suddenly speak about Gibbs? Why not also a bit about MCMC on piecewise models first?

5- We do not speak about the innovation of using delta calculus to eliminate complicated determinism not studied so far.

GMs:

1- Do we need to describe Gibbs here as a standalone topic?

….. ………….. ………………………………. ……………………………………………………………………..

In applications of probabilistic inference, instead of direct observation of random variables (e.g., $X$ and $Y$), functions of them may be observed (e.g., $X\codtY = c$).

This leads to posterior distributions which are only positive on sub-manifolds of the parameter space. Carrying out inference on such models is by no means trivial \cite{pennec2006intrinsic }.

To evade the complications, state-of-the-art probabilistic inference tools suggest adding noise to the observation and performing MCMC based approximate inference. Nonetheless, this does not solve the problem: If the added noise is large then the approximation bounds can be arbitrarily large and if it is small, the sampling mixing rate can be arbitrarily slow.

The other potential solution is to reduce the dimensionality of the posterior distribution via Jacobian-based random variable transformations. Measure theoretic subtleties aside, such transformations are only applicable when the observed function is invertible with respect to at least one random variable. Using the properties of the Dirac delta, our first contribution is to propose a dimension reduction method that is more general in the sense that the observed function is not required to be invertible but should be solvable with one or several distinct roots. Up to our knowledge, this is the first time that Dirac delta is used in this context.

Nonetheless, dimension reduction (either carried out via Jacobian-based or Dirac delta-based mechanism) does not completely eliminate the problem since as it will be shown shortly, the produced low dimensional distributions are highly piecewise and multimodal. Exact inference on such models is almost never possible and the convergence rate of the approximate alternatives can be extremely low.

…

For instance, in the adjacency of discontinuities (borders of pieces) the performance of Hamiltonian Monte Carlo (HMC) can be severely low. The reason is that the leapfrog-based mechanism that HMC relies on (to simulate the Hamiltonian dynamics) relies on the assumption of smoothness that does not hold in these areas.

Slice sampling suffers from the multimodal nature of the distributions that are in the focus of the present work. Similarly, near the borders of partitions, the acceptance rate of Metropolis-Hastings(MH) is typically low since in such areas the difference (e.g.\ KL-divergence) between MH’s \emph{proposal distribution} and the suddenly varying target distribution is often significant.

The exception is Gibbs sampling. The latter method can be regarded as a particular variation of MH where the proposals are directly chosen from the target distribution and therefore follow the sudden changes and multimodalities.

…

Nonetheless, Gibbs samplers can be quite slow since the computation of univariate CDFs that Gibbs relies on can be quite costly and in general cannot be performed in closed form.

In the second part of the paper, we focus on a highly expressive family of piecewise models.

This family consists of real-space piecewise distributions where the partitioning hyperplanes (i.e.\ borders) are made of polynomials where the degree of each variable is at most 2 and the function value in each partition has a fractional form where the denominator is factorized into terms where the degree of each variable is at most 2.

We show that univariate integrals (w.r.t.\ any variable) of such piecewise functions can always be computed analytically. In this context, second contribution of the paper is a variation of Gibbs sampling (namely, \emph{symbolic Gibbs}).

The idea is to construct a mapping from variables to their associated univariate CDFs that are computed symbolically and prior to the sampling. By this trick, the actual CDF computation per sample and variable is reduced instantiation of the associated symbolic CDF. This saves a tremendous amount of redundant computations and leads to a sampler that is at least an order of magnitude faster than the baseline.

Note that the reference family of models is rich enough to approximate arbitrary distributions up to arbitrary precision. It should also be pointed out that symbolic Gibbs addresses the problem of sampling from piecewise distributions in general. Sampling from the reduced dimension distributions is only an application of the former problem. Nonetheless, since it is the main motivation and part of our first contribution, we do not study other applications of symbolic Gibbs sampling throughout.

….

Our second contribution is Gibbs on the class labla.

The class is reach enough because…

Our solution is applicable to any piecewise but due to its importance, we concentrate on determinism throughout.

///////////////////////////////////////

@article{pennec2006intrinsic,

title={Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements},

author={Pennec, Xavier},

journal={Journal of Mathematical Imaging and Vision},

volume={25},

number={1},

pages={127--154},

year={2006},

publisher={Springer}

}