Randomness and Computation
Joint Workshop "New Horizons in Computing" and "Statistical Mechanical Approach to Probabilistic Information Processing" (18-21 July, 2005, Sendai, Japan)

CFTP algorithms for perfect sampling

Tomomi MATSUI [1] and Shuji KIJIMA [2]
Graduate School of Information Science and Technology,
University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan.

1 Introduction

Markov Chain Monte Carlo (MCMC) methods are used to sample from complicated multivariate distributions with normalizing constants that may not be computable in practice and from which direct sampling is not feasible. This is typically the case in image analysis and spatial statistics. Asisted by methods for convergence diagnostis, the MCMC user has to determine how long to run the simulation. However, the method always produces an approximate sample.

A *perfect sampler* is a method which produces a sample exactly according to the target distribution. Propp and Wilson [5] devised a perfect sampler, called *Coupling From The Past* (CFTP), based on an ergodic Markov chain which produces samples exactly according to a unique stationary distribution. The CFTP algorithm automatically determines how long it needs to run.

This talk gives a review of the CFTP algorithms. The concepts of coupling and monotone CFTP are introduced.

2 Coupling From the Past Algorithm

When we simulate an ergodic Markov chain for infinite time, we can gain a sample exactly according to the stationary distribution. Suppose that there exists a chain from infinite past, then a possible state at the present time of the chain for which we can have an evidence of the uniqueness without respect to an initial state of the chain, is a realization of a random sample exactly from the stationary distribution. This is the key idea of CFTP.

Suppose that we have an ergodic Markov chain MC with finite state space Ω and transition matrix P. The transition rule of the Markov chain $X \mapsto X'$ can be described by a deterministic function $\phi : \Omega \times [0,1) \to \Omega$, called *update function*, as follows. Given a random number Λ uniformly distributed over $[0,1)$, update function ϕ satisfies that $\Pr(\phi(x,\Lambda) = y) = P(x,y)$ for any $x, y \in \Omega$. We can realize the Markov chain by setting $X' = \phi(X, \Lambda)$. Clearly, update function corresponding to the given transition matrix P is not unique. The result of transitions of the chain from the time t_1 to t_2 ($t_1 < t_2$) with a sequence of random numbers $\boldsymbol{\lambda} = (\lambda[t_1], \lambda[t_1+1], \ldots, \lambda[t_2-1]) \in [0,1)^{t_2-t_1}$ is denoted by $\Phi_{t_1}^{t_2}(x, \boldsymbol{\lambda}) : \Omega \times [0,1)^{t_2-t_1} \to \Omega$ where $\Phi_{t_1}^{t_2}(x, \boldsymbol{\lambda}) \stackrel{\text{def.}}{=} \phi(\phi(\cdots(\phi(x, \lambda[t_1]), \ldots, \lambda[t_2-2]), \lambda[t_2-1])$. We say that a sequence $\lambda \in [0,1)^{|T|}$ satisfies the *coalescence condition*, when $\exists y \in \Omega$, $\forall x \in \Omega$, $y = \Phi_T^0(x, \boldsymbol{\lambda})$.

With these preparation, standard Coupling From The Past algorithm is expressed as follows.

CFTP Algorithm [5]

Step 1. Set the starting time period $T := -1$ to go back, and set $\boldsymbol{\lambda}$ be the empty sequence.

Step 2. Generate random real numbers $\lambda[T], \lambda[T+1], \ldots, \lambda[\lceil T/2 \rceil - 1] \in [0,1)$, and insert them to the head of $\boldsymbol{\lambda}$ in order, i.e., put $\boldsymbol{\lambda} := (\lambda[T], \lambda[T+1], \ldots, \lambda[-1])$.

Step 3. Start a chain from all elements $x \in \Omega$ at time period T, and run each chain to time period 0 according to the update function ϕ with the sequence of numbers in $\boldsymbol{\lambda}$. (Note that every chain uses the common sequence $\boldsymbol{\lambda}$.)

[1] E-mail: tomomi@mist.i.u-tokyo.ac.jp
[2] E-mail: kijima@misojiro.t.u-tokyo.ac.jp

Step 4. [Coalescence check]

If $\exists y \in \Omega$, $\forall x \in \Omega$, $y = \Phi_T^0(x, \boldsymbol{\lambda})$, then return y and stop.

Else, update the starting time period $T := 2T$, and go to Step 2.

Theorem 1 (CFTP Theorem [5]) *Let MC be an ergodic finite Markov chain with state space Ω, defined by an update function $\phi : \Omega \times [0, 1) \to \Omega$. If the CFTP algorithm terminates with probability 1, then the obtained value is a realization of a random variable exactly distributed according to the stationary distribution.*

Theorem 1 gives a (probabilistically) finite time algorithm for infinite time simulation. However, simulations from all states executed in Step 3 is a hard requirement.

Suppose that there exists a partial order "$\succeq$" on the set of states Ω. A transition rule expressed by a deterministic update function ϕ is called *monotone* (with respect to "$\succeq$") if $\forall \lambda \in [0, 1)$, $\forall x, \forall y \in \Omega$, $x \succeq y \Rightarrow \phi(x, \lambda) \succeq \phi(y, \lambda)$. For ease, we also say that a chain is *monotone* if the chain has a *monotone* transition rule.

Theorem 2 (monotone CFTP [5, 1]) *Suppose that a Markov chain defined by an update function ϕ is monotone with respect to a partially ordered set of states $(\Omega, \succeq)$, and $\exists x_{\max}, \exists x_{\min} \in \Omega$, $\forall x \in \Omega$, $x_{\max} \succeq x \succeq x_{\min}$. Then the CFTP algorithm terminates with probability 1, and a sequence $\lambda \in [0, 1)^{|T|}$ satisfies the coalescence condition, i.e., $\exists y \in \Omega$, $\forall x \in \Omega$, $y = \Phi_T^0(x, \boldsymbol{\lambda})$, if and only if $\Phi_T^0(x_{\max}, \boldsymbol{\lambda}) = \Phi_T^0(x_{\min}, \boldsymbol{\lambda})$.*

When the given Markov chain satisfies the conditions of Theorem 2, we can modify CFTP algorithm by substituting Step 4 (a) by

Step 4. (a)′ If $\exists y \in \Omega$, $y = \Phi_T^0(x_{\max}, \boldsymbol{\lambda}) = \Phi_T^0(x_{\min}, \boldsymbol{\lambda})$, then return y and stop.

The algorithm obtained by the above modification is called a monotone CFTP algorithm.

3 Discussion

Perfect simulation by CFTP and other methods is currently a thriving research topic. The research and development of perfect sampler following the paper by Propp and Wilson [5] have been extensive. Fill [2] proposed an interruptible method for perfect sampling based on a Makov chain. Wilson [7] devised a read onces algorithm which reduces the memory space required in an ordinary (monotone) CFTP algorithm. Murdoch and Green [4] proposed a perfect sampling method for continuous distributions. See the survey [1] and the textbook [3] also.

A perfect sampling algorithm is a remarkable tool, freeing users from the problems of convergence assesment connected to standard MCMC. New and effective perfect sampling algorithm for more general models will be developed through ongoing and future research.

References

[1] X. K. Dimakos: "A guide to exact simulation," International Statistical Review, **69** (2001), 27–48.

[2] J. Fill: "An interruptible algorithm for perfect sampling via Markov chains," The Annals of Applied Probability, **8** (1998), 131–162.

[3] O. Häggström: *Finite Markov Chains and Algorithmic Application*, London Mathematical Society, Student Texts **52**, Cambridge University Press, 2002.

[4] D. J. Murdoch and P. J. Green: "Exact sampling from a continuous state space," Scandinavian Journal of Statistics, **25** (1998), 483–502.

[5] J. Propp and D. Wilson: "Exact sampling with coupled Markov chains and applications to statistical mechanics," Random Structures and Algorithms, **9** (1996), 232–252.

[6] J. Propp and D. Wilson: "How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph," Journal of Algorithms, **27** (1998), 170–217.

[7] D. Wilson: "How to couple from the past using a read-once source of randomness," Random Structures and Algorithms, **16** (2000), 85–113.