The Efficiency Study of Splitting and Branching in Monte Carlo Method

Ilya N. Medvedev

Institute of Computational Mathematics
and Mathematical Geophysics
(Siberian Branch of the Russian Academy of Sciences)
prospect Akademika Lavrentjeva, 6,
Novosibirsk, 630090, Russia
e-mail: min@osmf.sscc.ru
web: www.sscc.ru

2011

CONTENTS

1. Integral equation, weight “collision estimator”,
2. Computational efficiency (computation cost) under trajectory branching
3. Computational efficiency (computation cost) under trajectory splitting
    - At first time in Markov chain (=1)
    - At some time in Markov chain (>1)

$$\varphi^*(x) = \int\limits_X k(x,x')\varphi^*(x')\,dx' + h(x) \ \text{ or } \ \varphi^* = K^*\varphi^* + h, \tag{1}$$

$$\int k(x',x)\,dx = q(x') \leq 1-\delta,\ \delta > 0, \ \ X \ is \ \ bounded \ \ domain \ \ in \ \ R^n.$$

$$\varphi^*, h \geq 0, \ \ \varphi^*, h \in L_\infty(X), \ \ K^* \in [L_\infty(X) \to L_\infty(X)],$$

Weighted “collision estimator”

$$x_1, \ldots, x_N$$

$$p(x,x') - \text{simulated transition distribution density } \ x \to x'$$

$$Q_0(x_0) = 1, \quad Q_n = Q_{n-1}\frac{k(x_{n-1},x_n)}{p(x_{n-1},x_n)}, \ \ \xi_x = h(x) + \sum_{n=1}^{N} Q_n h(x_n) \tag{2}$$

$$\varphi^*(x) = \mathrm{E}\xi_x = h(x) + \mathrm{E}\sum_{n=1}^{N} Q_n h(x_n), \ \ x_0 \equiv x, \tag{3}$$

The method of recurrent probabilistic averaging [1]

$$\varphi^* = \sum_{n=0}^{\infty} K^{*n}h, \quad \xi_{x_0} = h(x_0) + \sum_{n=1}^{N} Q_n h(x_n) = \sum_{n=0}^{\infty} \Delta_n Q_n h(x_n),$$

where $Q_0 \equiv 1$, $Q_n = Q_{n-1}q(x_{n-1}, x_n)$;

$$q(x_{n-1}, x_n) = \frac{k(x_{n-1}, x_n)}{p(x_{n-1}, x_n)}$$

Δ_n – indicator of “non-break” till x_n. Under presented assumptions: $\mathrm{E}\xi_{x_0} = \varphi^*(x_0)$ [3, 4].

$$\xi_{x_0} = h(x_0) + \Delta_1 q(x_0, x_1)\xi_{x_1}. \tag{4}$$

Recurrence $\xi_x = h(x) + \delta_x q(x, x')\xi_{x'}$ defines weight estimator uniquely. Here δ_x – indicator of “non-break” in the transition $x \rightarrow x'$, and $\delta_{x_0} = \Delta_1$. If $q(x, x') \leq 1$ then variance $\mathrm{D}\xi_{x_0}$ is knowingly finite [3, 4]. In other case $\mathrm{D}\xi_{x_0}$ can be infinite.

Let us consider the *number of the branches* $\nu(x,x')$ and *probability* $\alpha(x,x')$ so

$$P(\nu(x,x') = r(x,x')) = 1 - \alpha(x,x'),$$
$$P(\nu(x,x') = r(x,x') + 1) = \alpha(x,x').$$

Let ζ_x is defined by the recurrence:

$$\zeta_x = h(x) + \delta_x \frac{q(x,x')}{\mathrm{E}\nu(x,x')} \sum_{i=1}^{\nu(x,x')} \zeta_{x'}^{(i)}, \qquad (5)$$

where $\{\zeta_{x'}^{(i)}\}$ – independent realizations of $\zeta_{x'}$.

Lemma 1 *Under defined assumptions:* $\mathrm{E}\zeta_{x_0} = \varphi^*(x_0)$.

Proof Using the Wald identity we have:

$$\mathrm{E} \sum_{i=1}^{\nu(x_0,x_1)} \zeta_{x_1}^{(i)} = \mathrm{E}\nu(x,x')\mathrm{E}\zeta_{x_1}.$$

$$\mathrm{E}\zeta_x = \int_X k(x,x')\mathrm{E}\zeta_{x'}dx' + h(x).$$

THEOREM 1 *The function* $\mathrm{E}\zeta_x^2$ *is defined [1] by Neumann series for equation*

$$\mathrm{E}\zeta_x^2 = \int\limits_X k(x,x')\frac{q(x,x')}{\mathrm{E}\nu(x,x')}\mathrm{E}\zeta_{x'}^2 dx' + H(x), \tag{6}$$

$$H(x) = h(x)[2\varphi^*(x) - h(x)] + \int\limits_X k(x,x')\psi(x,x')\varphi^{*2}(x')dx',$$

$$\psi(x,x') = \frac{q(x,x')\mathrm{E}\{\nu(x,x')(\nu(x,x')-1)\}}{(\mathrm{E}\nu(x,x'))^2}.$$

Direct simulation($q(x,x') \equiv 1$) without branching $\xi_x = h(x) + \delta_x \xi_{x'}$. Then [2]

$$\mathrm{E}\xi_x^2 = \int\limits_X k(x,x')\mathrm{E}\xi_{x'}^2 dx' + h(x)[2\varphi^*(x) - h(x)]$$

$$If \ \ \mathrm{E}\nu(x,x') = q(x,x') \ \ then \ \ \mathrm{E}\zeta_x^2 \geq \mathrm{E}\xi_x^2 \ \ or \ \ \mathrm{D}\zeta_x^2 \geq \mathrm{D}\xi_x^2.$$

Computational efficiency (computation cost)

Let us define **computation cost** as

$$S = T\mathrm{D}\xi$$

the product of the average time T needed for one realization (trajectory) of the ξ and the variance estimator $\mathrm{D}\xi$.

Let us suppose that the average simulation time t_p ($p(x,x')$) approximately equals to average simulation time t_k $(k(x,x'))$ for the one transition $x \to x'$ and $T_{(\cdot)} = t_{(\cdot)}\mathrm{E}N_{(\cdot)}$ where $\mathrm{E}N_{(\cdot)}$ is the expectation of the number of the state at which the trajectory terminates (under use of the simulation density $(\cdot)$).

$$n(x) = \int\limits_X k(x,x')n(x')\,dx' + I_{\{x\in X\}},$$

$$\tilde{\zeta}_x = I_{\{x\in X\}} + \delta_x \sum_{i=1}^{\nu(x,x')} \tilde{\zeta}_{x'}^{(i)}, \;\; \tilde{\xi}_x = I_{\{x\in X\}} + \delta_x \tilde{\xi}_{x'}, \;\; \mathrm{E}\tilde{\zeta}_x = \mathrm{E}N_{(p)}(x) = n(x) = \mathrm{E}N_{(k)}(x) = \mathrm{E}\tilde{\xi}$$

$$S_b(x) = \mathrm{D}\zeta_x t_b \mathrm{E}N_{(p)}(x) \geq S_d(x) = \mathrm{D}\xi_x t_d \mathrm{E}N_{(k)}(x)$$

Particle fission coefficient The problem of particle fission coefficient computation in convex domain D with absorber outside. $x_0, ..., x_N$ - the chain of particle collisions with the elements of D. After each collision with given probabilities we have scattering, absorption or fission with average ν particles.

Let $r_0 \in \mathrm{D}$ be initial particle position with initial direction ω_0. The original problem is closely connected with computation of the average number $\varphi^*(r_0, \omega_0)$ of the take-off particles where

$\varphi^* = K^*\varphi^* + h$ and $h(r, \omega) = 1$ for $r \notin \mathrm{D}$ and $h(r, \omega) = 0$ otherwise.

It is well-known that there exists such ν^* so if $\nu > \nu^*$ then the process is above-critical ($\varphi^*(x_0) = +\infty$). If $\nu < \nu^*$ then the environment is subcritical ($\varphi^*(x_0) < +\infty$). The simple weight algorithm for estimating $\varphi^*(x_0)$ is to simulate the next fission as scattering and multiply the particle weight by ν. In this case $\mathrm{E}\xi_x^2 = g < \infty$ where

$$g = K_p^* g + h[2\varphi^* - h], \qquad K_p^* \rightarrow \quad k^2(x, x')/p(x, x') \;\; if \;\; \nu^2 < \nu^*. \tag{7}$$

It is more natural to use the branching algorithm with $[\nu]$ и $[\nu] + 1$ particles fussed at the collision point. Due to the Theorem 1 we have $\mathrm{E}\zeta_x^2 < +\infty$ при $\nu < \nu^*$.

The result

$$\varphi^*(x) < C < +\infty \; \forall x, \;\; then \;\; \mathrm{D}\zeta_x < \infty$$

Condition $\rho(K^*) < 1$ is not required.

Splitting

Following [3], let us introduce the notation: $\zeta = g(\lambda, \eta)$ with j.d.f. $f(x,y)$ and

$$\mathrm{E}\zeta = \int\limits_X f(x,y)g(x,y)dxdy = \int\limits_X f_1(x)\mathrm{E}[\zeta|x]dx,$$

$$\mathrm{E}[\zeta|x] = \int\limits_X f_2(y|x)g(x,y)dy, \quad f_1(x) = \int_Y f(x,y)dy$$

where $f_1(x)$ is the density of the absolute distribution of λ; $f_2(y|x)$ is the density of the conditional distribution of η when $\lambda = x$. $\mathrm{DE}[\zeta|\lambda] \leq \mathrm{D}\zeta = \mathrm{DE}[\zeta|\lambda] + \mathrm{ED}[\zeta|\lambda]$

Let $\lambda \sim f_1(x)$ and let $n \geq 1$ be an integer. G.A. Mikhailov [3] used the following estimate:

$$\zeta_n = \frac{1}{n}\sum_{i=1}^{n} g(\lambda, \eta_i), \quad \mathrm{E}\zeta_n = I, \quad \mathrm{D}\zeta_n = \mathrm{DE}[\zeta|\lambda] + \frac{\mathrm{ED}[\zeta|\lambda]}{n} = A_1 + \frac{A_2}{n}$$

and demonstrates that

$$n^* = \sqrt{\frac{A_2 t_1}{A_1 t_2}} \;\; \textit{minimizes} \;\; S_n = (t_1 + nt_2)(A_1 + \frac{A_2}{n}) \leq S_0 = (t_1 + t_2)(A_1 + A_2)$$

where t_1 average computation time for λ and t_2 av. comp. time for η $(\lambda = x)$.

Splitting

Let us define the direct simulation of the collision estimator with splitting the trajectory only at second transition $x_1 \to x_2$

$$\zeta_x = h(x) + \delta_x \zeta_{x'}, \quad \zeta_{x'} = h(x') + \delta_{x'} \frac{1}{\nu} \sum_{i=1}^{\nu} \xi_{x''}^{(i)}.$$

$$\mathrm{E}\zeta_x^2 = (A_h) + (A_\nu) = \left(h(x)[2\varphi^*(x) - h(x)] + \int\limits_X k(x,x')h(x')[2\varphi^*(x') - h(x')]dx' \right) +$$

$$\left(\int\limits_X k(x,x') \int\limits_X k(x',x'') \frac{\nu - 1}{\nu} \varphi^{*2}(x'') dx'' dx' + \int\limits_X k(x,x') \int\limits_X \frac{1}{\nu} k(x',x'') \mathrm{E}\xi_{x''}^2 dx'' dx' \right)$$

$$\mathrm{E}\xi_x^2 = (A_h) + (A_d) = (A_h) + \int\limits_X k(x,x') \int\limits_X k(x',x'') \mathrm{E}\xi_{x''}^2 dx'' dx'$$

Let us compare the **computation cost**

$$S_\nu(x) = (t_1 + \nu t_2)(A_h + A_\nu) \quad with \quad S_1(x) = (t_1 + t_2)(A_h + A_d)$$

where t_1 average computation time for $x_0 \to x_2$ and t_2 av. comp. time for $x_2 \to x_N$.

It is easy to check that

$$(S_\nu(x) - S_1(x))'_\nu = \frac{t_1}{\nu^2}\big([K^*\varphi^{*2}](x) - A_d\big) + t_2\big(A_h + [K^*\varphi^{*2}](x) - \varphi^{*2}(x)\big)$$

and

$$(S_\nu(x) - S_1(x))'_\nu = 0 \;\; if \;\; \nu^* = \sqrt{\frac{t_1(A_d - [K^*\varphi^{*2}](x))}{t_2(A_h + [K^*\varphi^{*2}](x) - \varphi^{*2}(x)\big)}}.$$

Since

$$(S_\nu(x) - S_1(x))''_{\nu\nu}(\nu^*) > 0 \;\; then \;\; \nu^* \;\; provides \;\; minimum \;\; to \;\; S_\nu(x) - S_1(x)$$

and

$$S_{\nu^*}(x) - S_1(x) \geq 0 \;\; \forall x \in X$$

Let us note that in general A_h and A_d can be estimated by the result from special a priori calculations.

Splitting Let us consider the direct simulation with splitting the trajectory only at first transition $x_0 \to x_1$. In this case

$$\zeta_x = h(x) + \delta_x \frac{1}{\nu} \sum_{i=1}^{\nu} \xi_{x'}^{(i)}, \quad .$$

$$\mathrm{E}\zeta_x^2 = \int\limits_X \frac{1}{\nu}\, k(x,x')\mathrm{E}\xi_{x'}^2 dx' + h(x)[2\varphi(x)^* - h(x)] + \int\limits_X k(x,x')\frac{\nu-1}{\nu}\varphi^{*2}(x')dx'. \quad (8)$$

Computation cost $\quad S_\nu(x) - S_d(x) = \nu t_k \mathrm{D}\zeta_x - t_k \mathrm{D}\xi_x =$

$$t_k\Big(\nu\mathrm{D}\zeta_x - \int\limits_X k(x,x')\mathrm{E}\xi_{x'}^2 dx' - h(x)[2\varphi^*(x) - h(x)] + \varphi^{*2}(x)\Big) =$$

$$\nu t_k(\nu-1)\Big(h(x)[2\varphi^*(x) - h(x)] + \int\limits_X k(x,x')\varphi^{*2}(x')dx' - \varphi^{*2}(x)\Big) =$$

$$\nu t_k(\nu-1)\big(\varphi^{*2}(x) - [K^*\varphi^*]^2(x) + [K^*\varphi^{*2}](x) - \varphi^{*2}(x)\big) =$$

$$\nu t_k(\nu-1)\big([K^*\varphi^{*2}](x) - [K^*\varphi^*]^2(x)\big) \geq 0.)$$

((Cauchy-Bunyakovsky) Schwarz inequality)

Список литературы

[1] Medvedev I.N., Mikhailov G.A. Recurrent partial averaging in the theory of weighted Monte Carlo methods // Russ. J. Numer. Anal. Math. Modelling. - 2009 - Vol. 24, є 3, pp.261-277

[2] S.M. Ermakov and G.A. Mikhailov, *Statistical Simulation.* Nauka, Moscow, 1982 (in Russian).

[3] *Mikhailov G.A.* Optimization of Monte-Carlo Weighted Methods. Springer-Verlag, 1992.

[4] *Mikhailov G.A., Voytishek A., V* Computational statistical modelling. Monte-Carlo method. M.: "Akademia 2006.

[5] *Mikhailov G.A.* New Monte Carlo Methods with Estimating Derivatives. VSP, Utrecht, The Netherlands, 1995.