

УДК 519.71 + 519.24
ББК 22.1

О СРАВНЕНИИ СЛУЧАЙНЫХ БЛУЖДАНИЙ¹

Грязина Е.Н.²

*(Учреждение Российской академии наук Институт проблем
управления им. В.А. Трапезникова РАН, Москва)*

В работе сравниваются три метода генерации точек в ограниченной многомерной области. Все три метода – Hit-and-Run, Barrier Monte Carlo, Billiard Walk – представляют собой случайные блуждания по области и порождают асимптотически равномерное распределение точек. Обсуждаются критерии сравнения таких генераторов, предлагается новый критерий сравнения на основе коэффициента Джини, и, наконец, приводятся результаты сравнения методов в тестовых областях.

Ключевые слова: рандомизированные алгоритмы, метод Монте-Карло, марковские цепи, случайные блуждания, проверка статистических гипотез, коэффициент Джини.

1. Методы генерации точек

Генерирование точек в ограниченной области со сложной геометрией востребовано во многих задачах моделирования систем, анализа и синтеза в задачах с неопределенностью [14, 12, 9, 2].

Хронологически первым является метод Hit-and-Run [15, 13]. С минимальными требованиями к области Hit-and-Run легок в применении и эффективен даже для несвязных областей [11]. Однако, чудовищные оценки для скорости достижения рав-

¹ Работа выполнена при финансовой поддержке РФФИ, грант № 14-07-00067 А. Автор признательна Б.Т. Поляку за ценное обсуждение содержания статьи.

² Грязина Елена Николаевна, кандидат физико-математических наук (Москва, ул. Профсоюзная, д. 65).

номерного распределения [6] и неудовлетворительное поведение метода на практике побудило предложить другие алгоритмы. Попытки использовать идеи метода в внутренней точки [7], а именно, барьерные функции и вписанные эллипсоиды, привели к методу Barrier Monte Carlo [10]. Этот метод существенно превосходит Hit-and-Run в скорости получения приемлемо равномерных точек, но вместе с тем более сложен вычислительно. Наконец, недавний метод Billiard Walk [3] продемонстрировал высокую эффективность на тестовых задачах. В основу метода легли идеи Gas Kinetic Point Generator [5].

2. Способы сравнения

В данной работе мы систематически сравниваем вышеупомянутые методы по ряду критериев. В некоторых случаях достаточно визуально оценить полученные точки в проекции на показательную двумерную плоскость. Чаще по визуальной картинке судить о равномерности весьма затруднительно, тогда мы сравниваем функции распределения специально построенных случайных величин.

К полученным точкам можно применять весь арсенал методов проверки статистических гипотез, в частности, критерий Колмогорова-Смирнова [4] и критерий χ^2 [8]. Как правило, подобные тесты дают отрицательный результат, ведь при небольшом количестве точек их распределение далеко от равномерного. Поэтому на основе коэффициента Джини [1] предлагается новый критерий, который хорошо подходит для выборок с разным количеством точек.

3. Выводы

В приложениях при необходимости генерировать точки в многомерных областях со сложной геометрией достичь гарантированно равномерное распределение точек становится вычислительно неразрешимой задачей. Рассмотренные способы дают инженеру богатый арсенал средств для мотивированного выбора

алгоритмов генерации случайных точек.

Литература

1. GINI C. *Concentration and dependency ratios* (in Italian). - 1909. English translation in Rivista di Politica Economica. - 1997. - Vol. 87. - P. 769–789.
2. GORNOV A., FINKELSTEIN E. *Practical approximation algorithms for reachable set of dynamic controlled system* // IV International conference «Optimization and applications» (OPTIMA-2013). - Moscow. - 2013. – P. 74.
3. GRYAZINA E.N., POLYAK B.T. *Random sampling: Billiard Walk algorithm*//European Journal of Operational Research. – 2014. –Vol. 238. – P. 497-504.
4. KOLMOGOROFF A.N. *Sulla determinazione empirica di una legge di distribuzione* // Giornale dell' Istituto Italiano degli Attuari. 1933. – Vol. 4. – № 1. – P. 83-91.
5. LEE L., POOLLA H. *Statistical validation for uncertainty models* //Feedback, Control, Nonlinear Systems, and Complexity. - 1995. - Springer. - P. 131-149.
6. LOVASZ L., VEMPALA S. *Hit-and-run from a corner* //Proceedings of the 36th annual ACM symposium on Theory of computing, Chicago, IL, USA. - 2004. - P. 310–314.
7. NESTEROV Y., NEMIROVSKY A. *Interior Point Polynomial Methods in Convex Programming*. - 1994. - SIAM, Philadelphia.
8. PEARSON, K. *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling* // Philosophical Magazine Series 5 50. - 1900. - Vol. 302. - P. 157–175.
9. POLYAK, B. AND GRYAZINA E. *Robust stabilization via Hit-and-Run techniques* // IEEE Multi-Conference on Systems and Control. - 2009. - P. 537-541.
10. POLYAK, B. AND GRYAZINA E. *Markov Chain Monte Carlo method exploiting barrier functions with applications*

- to control and optimization* // IEEE Multi-Conference on Systems and Control. - 2010. - P. 1553–1557.
11. POLYAK, B. AND GRYAZINA E. *Randomized methods based on new Monte Carlo schemes for control and optimization* // Annals of Operational Research. - 2011. - Vol. 189. - № 1. - P. 343–356.
 12. RUBINSTEIN R., KROESE D. *Simulation and the Monte Carlo Method*. - 2008. - Wiley, NJ.
 13. SMITH R. *Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions* // Operations Research. - 1984. - Vol. 32. - № 6. - P. 1296–1308.
 14. TEMPO R., CALAFIORE G., DABBENE F. *Randomized Algorithms for Analysis and Control of Uncertain Systems, Communications and Control Engineering Series*. - 2004. - Springer-Verlag, London.
 15. TURCHIN V. *On the computation of multidimensional integrals by the Monte Carlo method* // Theory of Probability and its Applications. - 1971. - Vol. 16. - № 4. - P. 720–724.

ON THE COMPARISON OF RANDOM WALKS

Elena Gryazina, Institute for Control Sciences of RAS, Moscow,
Candidate of Science (gryazina@gmail.com).

Abstract: We compare three methods for sampling in a bounded multi-dimensional domain. All the methods – Hit-and-Run, Barrier Monte Carlo, Billiard Walk – are random walks and they generate asymptotically uniformly distributed samples. We discuss how to compare between these generators and propose new criteria based on Gini coefficient. Finally, we present comparison results for test domains.

Keywords: randomized algorithms, Markov Chain Monte Carlo, random walks, statistical hypothesis testing, Gini coefficient.