STATE RESEARCH CENTER OF RUSSIA
INSTITUTE FOR HIGH ENERGY PHYSICS

IHEP 2010-7

I.A. Kachaev[1]

Method to Obtain Better Accuracy for Monte-Carlo Integration over Phase Space with Acceptance

[1] E-mail: Igor.Katchaev@ihep.ru

Protvino 2010

UDK 519.25.256 M–24

Аннотация

Качаев И.А. Метод повышения точности Монте-Карло интегрирования по фазовому объему с аксептансом: Препринт ИФВЭ 2010-7. – Протвино, 2010. – 3 с., библиогр.: 4.

Предлагается метод ускорения сходимости Монте-Карло интегрирования при вычислении интегралов по фазовому объему, в которые входит аксептанс установки. Точность оценки интеграла $\int g(x)\varepsilon(x)\,dx$ методом Монте-Карло, где $g(x)$ — произвольная функция, а $\varepsilon(x)$ — аксептанс установки, может быть улучшена, если доступна хорошая оценка интеграла $\int g(x)\,dx$.

Abstract

Kachaev I.A. Method to Obtain Better Monte-Carlo Accuracy for Phase Space Integration with Acceptance: IHEP Preprint 2010-7. – Protvino, 2010. – p. 3, refs.: 4.

Method is suggested to obtain better Monte-Carlo accuracy for phase space integration with acceptance. Better Monte-Carlo estimate can be obtained for the integral $\int g(x)\varepsilon(x)\,dx$ where $g(x)$ is arbitrary function and $\varepsilon(x)$ is setup acceptance if good estimate for the integral $\int g(x)\,dx$ is available.

1. Subtraction of the largest part in numerical integration

Often in high energy physics one need to describe a situation where particles born in reaction with given kinematics an matrix element pass through an experimental setup with given acceptance. In this situation integrals of the form

$$I = \int g(x)\,dx\,, \quad I_\varepsilon = \int g(x)\varepsilon(x)\,dx \tag{1}$$

are common say for partial wave analysis method. Here integration variable x includes both kinematical variables and variables which describes experimental setup, starting from interaction vertex up to hits in detectors. After proper coordinate transformation one can assume that integration here is over unit cube of significant dimension. Function $g(x)$ includes matrix element and kinematic factors, function $\varepsilon(x)$ describes acceptance of the setup.

Due to huge complexity such integrals are generally calculated by Monte-Carlo method, namely by generation of events with proper probability density and simulation of the passage of them through experimental setup. Convergence rate for Monte-Carlo integration is slow. This article suggests a method to obtain better convergence for the integrals like I_ε in (1).

It is known that Monte-Carlo mean square integration error is governed by the dispersion of integrated function. Let us introduce for clarity standard operators of mean value and dispersion for given function $f(x)$

$$E(f) \equiv \overline{f} \equiv \int f(x)\,dx\,, \quad D(f) \equiv E\left((f - E(f))^2\right) = \overline{f^2} - (\overline{f})^2 \tag{2}$$

Mean square integration error for integration of $f(x)$ over N points (generated events) is

$$\sigma(I) = \sqrt{D(f(x))/N} \tag{3}$$

One can see that integration error decreases when dispersion of the function to be integrated decreases. There are a lot of methods to decrease dispersion in the general case, they can be found in [1, 2].

Next idea is known as subtraction of the largest part. Let us find a function which is near to the function given but has a property that this new function can be integrated faster and/or with better accuracy than given one. In this case Monte-Carlo method can be used to integrate only a small difference between these functions. In our situation in most cases integral I is much simple to calculate that I_ε due to the fact that $g(x)$ does not depend on setup, has smaller dimension and I can often be calculated (semi)analytically.

In the case when a good estimate for I is available one can use a function $\varepsilon\, g(x)$ as a function near to $g(x)\varepsilon(x)$ to calculate I_ε. Here ε is a number which can be selected later to minimize dispersion of the difference between functions described above. For arbitrary ε one can write

$$\int \varepsilon(x)g(x)\,dx = \int \varepsilon\, g(x)\,dx + \int [\varepsilon(x)-\varepsilon]g(x)\,dx \tag{4}$$

We will calculate first term in the right part of (4) by some exact method, second term by the Monte-Carlo method. Since first term is assumed exact, calculation error will be from the second term only. Let I and I_ε be integrals estimate from pure Monte-Carlo. Let I_0 be exact value of integral without acceptance and I'_ε be improved estimate of integral with acceptance given by this method. In this notation we have from (4)

$$I'_\varepsilon = \varepsilon(I_0 - I) + I_\varepsilon \tag{5}$$

One should stress that according to (4) estimates of integrals I and I_ε here are not independent but obtained *from the same set of integration points (generated events)*. To find optimal ε let us write dispersion of integrated function for the second term in the right part of (4)

$$D(\varepsilon) = \overline{([\varepsilon(x)-\varepsilon]g(x))^2} - (\overline{[\varepsilon(x)-\varepsilon]g(x)})^2$$

One can find minimum of this value from condition $dD(\varepsilon)/d\,\varepsilon = 0$ using the fact that calculation of mean value and differentiation commutates. We use here the same symbol for dispersion $D(f(x))$ as functional and as a function of shift $D(\varepsilon)$, meaning is clean from the context. Optimal shift is

$$\varepsilon_0 = \frac{\overline{\varepsilon(x)g^2(x)} - \overline{\varepsilon(x)g(x)}\,\overline{g(x)}}{\overline{g^2(x)} - \overline{g(x)}\,\overline{g(x)}} \tag{6}$$

where denominator is $D(g(x))$. To estimate a gain in dispersion and so in integration error one can note that $D(\varepsilon)$ is quadratic in ε, namely

$$D(\varepsilon) = a\varepsilon^2 + b\varepsilon + c, \quad a = D(g(x)), \quad \varepsilon_0 = \frac{-b}{2a}, \quad D(0) - D(\varepsilon_0) = \frac{b^2}{4a}$$

We have a gain in dispersion which is always positive:

$$D(0) - D(\varepsilon_0) = \varepsilon_0^2\, D(g(x)) \tag{7}$$

Albeit ε_0 means some "mean acceptance" this value need not be positive. Example: integration domain is one dimensional, $0 < x < 1$,

$$\begin{array}{ll} \varepsilon(x) = 1, & g(x) = 0.5 \quad \text{if } x < 0.5 \\ \varepsilon(x) = 0, & g(x) = 1 \quad\;\; \text{if } x \geq 0.5 \end{array}$$

In this example $\varepsilon_0 = -1$ and $D(\varepsilon_0) = 0$ so integration with this method becomes exact. One should note that usage of *any* $\varepsilon > 0$ selected in advance gives in this example *worse* integration error than naive integration. Just this not absolutely correct method was used in [3].

2. Conclusions

This method has some advantages. First, with this method calculation error for I_ε depends only on nonuniformity of $\varepsilon(x)$ over phase space but not on its value. Next, due to the fact that $D(\varepsilon)$ is quadratic around ε_0 and integration error mildly depends on ε_0 estimate, value of ε_0 can be calculated with moderate degree of accuracy. Next, ε_0 can be calculated "from data" in parallel with calculation of integrals. To estimate I and I_ε with their errors via Monte-Carlo one need to keep sums of

$$g(x), \quad g^2(x), \quad \varepsilon(x)g(x), \quad \varepsilon^2(x)g^2(x)$$

To estimate in addition ε_0 one need to add to this list a sum of $\varepsilon(x)g^2(x)$ so additional calculations are minimal. After the loop over Monte-Carlo events one can calculate ε_0 from (6), integrals from (5) and finally error estimates from (7), (3).

Another advantage of this method is that it is "transparent" and can be combined with any other Monte-Carlo optimization: importance sampling, symmetrization, quasi Monte-Carlo, etc. The only requirement is that one need a good estimate of the integral of given function over phase space without acceptance.

This method was used in the investigation of the $\pi^+\pi^-\pi^-$ system with partial wave analysis method [4].

References

[1] James F, Monte Carlo Theory and Practice, Rep. Prog. Phys. Vol 43, 1980.

[2] Sobol I.M., Numerical methods for Monte-Carlo, Moscow, 1973 (in russian)

[3] J.D. Hansen et al., Nucl. Phys. **B81** (1974) 403

[4] D.I. Amelin et al., Phys. Lett. B 356 (1995), 595.

Received June 7, 2010.

Препринт отпечатан с оригинала-макета, подготовленного автором.

И.А. Качаев
Метод повышения точности Монте-Карло интегрирования по фазовому объему с аксептансом.

Оригинал-макет подготовлен с помощью системы **LaTeX.**

Подписано к печати 17.06.10. Формат $60 \times 84/16$. Офсетная печать.
Печ.л. 0,43. Уч.-изд.л. 0,48. Тираж 80. Заказ 54. Индекс 3649.

ГНЦ РФ Институт физики высоких энергий
142281, Протвино Московской обл.

Индекс 3649

П Р Е П Р И Н Т 2010–7, И Ф В Э, 2010