Лекция № 2

Цель работы

Изучение теории применения методов Монте-Карло.

Методические указания

Методы Монте-Карло – это общее название группы методов для решения различных задач с помощью *случайных последовательностей*. Исторически они возникли на базе выборочного метода в статистике и называются также методами *статистических испытаний*. Первоначально метод Монте-Карло использовался главным образом для решения задач математической физики, где традиционные численные методы оказались мало пригодными. Далее его влияние распространилось на теорию массового обслуживания, задачи теории игр и математической экономики, задачи теории передачи сообщений при наличии помех и ряд других.

Идея метода заключается в следующем. Вместо того чтобы описывать исследуемый случайный процесс аналитически, составляется алгоритм, *имитирующий* этот процесс. Как правило, программа составляется для осуществления одного случайного испытания. Затем это испытание повторяется N раз, причем каждый опыт не зависит от остальных, и результаты всех опытов усредняются.

Для применения методов Монте-Карло достаточно описания вероятностного процесса и не обязательна его формулировка в виде интегрального уравнения. Оценка погрешности метода чрезвычайно проста. Точность слабо зависит от размерности пространства.

Общий курс методов Монте-Карло

Важнейший прием построения методов Монте-Карло — сведение некой задачи к расчету математического ожидания: для т ого чтобы приближенно вычислить некоторую скалярную величину a, надо придумать такую случайную величину ξ, что $\mathbf{M}\xi = a$; тогда вычислив N независимых значений $\xi_1, \ldots, \xi_N$ этой величины ξ, можно считать, что

$$a \approx {}^{1}\!/_{N}\,(\xi_1 + \cdots + \xi_N).$$

Пример. Требуется оценить объем V_G некоторой ограниченной пространственной фигуры G.

Решение. Выберем параллелепипед $П$, содержащий G, объем которого $V_П$ известен (см. рис. _). Выберем N случайных точек, равномерно распределенных в $П$, и обозначим через N' количество точек, попавших в G. Если N велико, то очевидно, $N' : N \approx V_G : V_П$, откуда получаем оценку

$$V_G \approx V_G\, {}^{N'}\!/_{N}.$$

В этом примере случайная величина ξ равна $V_П$, если случайная точка попадает в G, и ξ равна нулю, если точка попадает в $П–G$. Нетрудно проверить, что математическое ожидание $\mathbf{M}\xi = V_G$, а среднее арифметическое

$$ {}^{1}\!/_{N}\,(\xi_1 + \cdots + \xi_N) = V_П\, {}^{N'}\!/_{N}.$$

Пример окончен.

Существует бесконечно много случайных величин ξ таких, что $\mathbf{M}\xi = a$. Теория методов Монте-Карло должна дать ответы на два вопроса:

1) Как выбрать удобную величину ξ для выполнения расчета той или иной задачи?
2) Как находить значения $\xi_1, \xi_2, \ldots$ произвольной случайной величины ξ?

Общий метод оценки математического ожидания

1. Сходимость метода. Чтобы оценить величину a, выберем N независимых реализаций $\xi_1, \ldots, \xi_N$ случайной величины ξ и вычислим среднее арифметическое

$$\bar{\xi}_N = {}^1/_N \sum_{i=1}^{N} \xi_i. \qquad (1)$$

Так последовательность одинаково распределенных независимых случайных величин, у которых существуют математическое ожидания, подчиняется закону больших чисел (теорема А.Я.Хинчина), то среднее арифметическое этих величин сходится по вероятности (стр. 34) к математическому ожиданию: при N→∞

$$\bar{\xi}_N \xrightarrow{P} a.$$

Таким образом, при больших N величина $\bar{\xi}_N \approx a$; и оценку (1) можно использовать когда существует $\mathbf{M}\xi = a$.

2. Погрешность метода. Предположим дополнительно, что случайная величина ξ имеет конечную дисперсию

$$\mathbf{D}\xi = \mathbf{M}(\xi^2) - (\mathbf{M}\xi)^2. \qquad (2)$$

Из курса теории вероятностей известно, что последовательность одинаково распределенных независимых случайных величин с конечными дисперсиями подчиняется центральной предельной теореме. Это означает, что для любых $x_1 < x_2$

$$\lim_{N\to\infty} P\left\{x_1 < {}^1\Big/_{\sqrt{N\mathbf{D}\xi}} \sum_{i=1}^{N} (\xi_i - a) < x_2\right\} = \left({}^1/_{\sqrt{2\pi}}\right) \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt.$$

При симметричном интервале x2=–x1=x получаем, что

$$\lim_{N\to\infty} P\left\{\left|{}^1/_N \sum_{i=1}^{N} (\xi_i - a)\right| < x\sqrt{{}^{\mathbf{D}\xi}/_N}\right\} = \Phi(x),$$

где $\Phi(x)$ — интеграл вероятностей (приводится в таблицах в справочниках):

$$\Phi(x) = \left({}^2/_{\sqrt{2\pi}}\right) \int_0^x e^{-\frac{t^2}{2}} dt.$$

Следовательно, при достаточно больших N

$$P\left\{\left|\bar{\xi}_N - a\right| < x\sqrt{{}^{\mathbf{D}\xi}/_N}\right\} \approx \Phi(x). \qquad (3)$$

Формула (3) содержит целое семейство оценок, зависящее от параметра x. Если задать любой коэффициент доверия β (стр. 31), то по таблице можно найти корень $x = x_\beta$ уравнения $\Phi(x) = \beta$. Тогда из (3) вытекает, что вероятность неравенства

$$\left|\bar{\xi}_N - a\right| < x_\beta\sqrt{{}^{\mathbf{D}\xi}/_N} \qquad (4)$$

приблизительно равна β. Чаще других используют коэффициент доверия $\beta = 0.997$ ($x_\beta = 3$) или $\beta = 0.95$ ($x_\beta = 1{,}96$). Значение соответствует так называемому «правилу трех сигм», где случайная величина $\bar{\xi}_N$ приближенно нормальна и ее среднее квадратичное уклонение $\sigma = \sqrt{{}^{\mathbf{D}\xi}/_N}$.

3. Вероятная ошибка метода. Иной подход к оценке ошибки связан с понятием вероятной ошибки

$$r_N = 0.6745\sqrt{\mathbf{D}\xi / N}. \quad (5)$$

Численный множитель 0.6745, это значение x_β, отвечающее $\beta = 0.50$. Название «вероятная ошибка» вызвано тем, что

$$P\{|\bar{\xi}_N - a| < r_N\} \approx 1/2 \approx P\{|\bar{\xi}_N - a| > r_N\},$$

т.е. одинаково вероятны ошибки, большие чем r_N, и ошибки, меньше чем r_N.

Величина r_N на практике часто используется для характеристики порядка ошибки: действительная ошибка $|\bar{\xi}_N - a|$ зависит от использованных в расчете случайных чисел и может оказаться в 2–3 раза больше, чем r_N, но может быть и меньше.

Используя r_N, мы оцениваем порядок ошибки, а используя (4) — верхнюю границу ошибки (с коэффициентом доверия β).

4. Эмпирическая оценка дисперсии. Когда мы приступаем к расчету $\mathbf{M}\xi$, значение дисперсии $\mathbf{D}\xi$, как правило, неизвестно. Хорошую теоретическую оценку для $\mathbf{D}\xi$ удается получить редко. Однако в большинстве задач величину $\mathbf{D}\xi$ нетрудно оценить эмпирически, в ходе расчетов a. Достаточно одновременно с вычислением $\Sigma\xi_i$ вычислять также и $\Sigma(\xi_i)^2$; так как при больших N

$$1/N \sum_{i=1}^{N} \xi_i^2 \approx \mathrm{M}(\xi^2),$$

то из (2) видно, что

$$\mathbf{D}\xi \approx 1/N \sum_{i=1}^{N} \xi_i^2 - \left[1/N \sum_{i=1}^{N} \xi_i\right]^2. \quad (6)$$

Формула (6) постоянно используется на практике. Доказано, что при небольших N более точна формула

$$\mathbf{D}\xi \approx \frac{1}{N-1}\sum_{i=1}^{N} \xi_i^2 - \frac{1}{N(N-1)}\left(\sum_{i=1}^{N} \xi_i\right)^2, \quad (7)$$

отличающаяся от прежней множителем $(1–1/N)$. Но в расчетах всегда $N>>10$, и разница между (7) и (6) невелика. К тому же $\mathbf{D}\xi$ используется только для оценки ошибки, поэтому погрешность порядка 10% в значении $\mathbf{D}\xi$ роли не играет.

5. Оценка ошибки без расчета дисперсии. Пусть дисперсия $\mathbf{D}\xi$ конечна. Оценку $\bar{\xi}_N - a$ погрешности можно получить и по-другому (если сложно организовать расчеты по 6).

Предположим, что $N=mN_1$, где m – небольшое натуральное число $m\geq 3$, а N_1 настолько велико, что распределение случайной величины

$$\zeta = 1/N_1 \sum_{i=1}^{N_1} \xi_i \quad (8)$$

можно считать близким к нормальному (по центральной предельной теореме). Очевидно, $\mathbf{M}\xi = a$.

Вместо вычисления $\bar{\xi}_N$, разделим задачу на m «вариантов» и вычислим m величин, которые можно считать независимыми реализациями ζ:

$$\zeta_1 = 1/N_1 \sum_{i=1}^{N_1} \xi_i,\ \zeta_2 = 1/N_2 \sum_{i=1}^{N_1} \xi_{i+N_1}, \ldots, \zeta_m = 1/N_1 \sum_{i=1}^{N_1} \xi_{i+N_1(m-1)}.$$

Используя теорему Фишера о подчинении закону распределении Стьюдента с $(m–1)$ степенью свободы, можно получить[1]: вероятность неравенства

[1] Написано с большим сокращением.

$$\left|\overline{\xi}_N - a\right| < t_{m-1,\beta}\sqrt{s^2/m-1} \quad (10)$$

Приблизительно равна β. Оценка (10) подобна оценке (4), но вместо неизвестной дисперсии сюда входит эмпирическая величина s2, которую можно легко вычислить по формуле (9). Но (4) применима при меньших N.

6. Случай $\mathbf{D}\xi=\infty$. Бесконечность дисперсии $\mathbf{D}\xi$ не препятствует приближению $\overline{\xi}_N$ к a. Порядок убывания при этом хуже $N^{-1/2}$. Обычно рекомендуется избегать методов расчета, в которых дисперсия осредняемой величины бесконечна.

Вычисление интегралов методом Монте-Карло

Для применения методов Монте-Карло достаточно описания вероятностного процесса и не обязательна его формулировка в виде интегрального уравнения.

Способ, основанный на истолковании интеграла как площади

Рассмотрим интеграл вида

$$I = \int_a^b \varphi(x)\,dx,$$

где подынтегральная функция неотрицательна и ограничена $0 \le \varphi(x) \le c$, исходя из истолкования интеграла как площади. Введем в рассмотрение двумерную случайную величину (X, Y), распределенную равномерно в прямоугольнике D с основанием $(b-a)$ и высотой c, плотность вероятности которой

$$f(x,y) = \begin{cases} \dfrac{1}{(b-a)c}, \text{для точек принадлежащих } D; \\ 0, \text{вне } D. \end{cases}$$

Составляющая X распределена в интервале (a, b) равномерно с плотностью распределения $1/(b-a)$; составляющая Y распределена в интервале $(0, c)$ с плотностью $1/c$. Если разыграно N точек (x_i, y_i), принадлежащих прямоугольнику D, из которых n точек оказались под кривой, то отношение площади, определяемой интегралом I, к площади прямоугольника D

$$\frac{\int_a^b \varphi(x)\,dx}{(b-a)c} \approx \frac{n}{N}.$$

Отсюда

$$\int_a^b \varphi(x)\,dx \approx (b-a)c\,\frac{n}{N}.$$

Тогда в качестве оценки интеграла можно взять:

$$I^* \approx (b-a)c\,\frac{n}{N}.$$

Вычисление интеграла способом усреднения подынтегральной функции

Рассмотрим интеграл вида

$$I = \int_a^b \varphi(x)\, dx.$$

Введем в рассмотрение случайную величину X, распределенную равномерно в интервале (a, b) с плотностью $f(x) = \frac{1}{(b-a)}$. Тогда математическое ожидание

$$\mathbf{M}[\varphi(x)] = \int_a^b \varphi(x) f(x)\, dx = \frac{1}{b-a}\int_a^b \varphi(x)\, dx.$$

Отсюда получаем

$$\int_a^b \varphi(x)\, dx = (b-a)\mathbf{M}[\varphi(x)].$$

Заменив математическое ожидание $\mathbf{M}[\varphi(x)]$ его оценкой – выборочной средней, получим оценку интеграла

$$I^* = (b-a)\frac{\sum_{i=1}^{N}\varphi(x_i)}{N},$$

где x_i — возможные значения случайной величины X, N — число испытаний. Так как случайная величина распределена равномерно в интервале (a, b) с плотностью $f(x) = {}^1/_{b-a}$, то x_i разыгрывается по формуле $x_i = a + (b-a)r_i$, где r_i — случайное число.

Важнейшие способы построения хороших оценок (способы уменьшения дисперсии)

Вероятная ошибка оценки (1) пропорциональна $\sqrt{\mathbf{D}\xi/_N}$. Скорость убывания ошибки невелика. Поэтому очень важно выбирать для расчета интегралов такие вычислительные схемы или, другими словами, такие случайные величины ξ, для которых дисперсия $\mathbf{D}\xi$ по возможности мала. Такие схемы называют способами уменьшения дисперсии, так как для таких способов дисперсия должна быть меньше дисперсии простейшего метода Монте-Карло.

1. Частичное аналитическое интегрирование. Если часть задачи можно решить аналитически, то обычно удается построить метод решения всей задачи с меньшей дисперсией. Но такие методы могут оказаться более трудоемкими и в конечном счете невыгодными.

1.1. Выделение главной части. Очевидный и весьма общий принцип: если главную часть задачи можно вычислить аналитически, то, как правило, выгодно считать методом Монте-Карло не всю задачу, а только «поправку» — разницу между всей задачей и главной частью. Уменьшение дисперсии может оказаться очень значительным.

1.2. Интегрирование по части области. Если известно аналитическое решение интеграла по некоторой части B от области G, то выгодно представить исходный интеграл в виде суммы

$$I = \int_{G_1} f(P)p(P)\, dP + C,$$

где G_1=G–B, а

$$C = \int_B f(P)p(P)\,dP,$$

1.2. Интегрирование по части переменных (понижение порядка интеграла). Если аналитически взять интеграл по некоторым из переменных, а по остальным переменным использовать метод Монте-Карло, то дисперсия уменьшится. Но нередко бывает, что после интегрирования по некоторым из переменных получаются более сложные формулы счета — дисперсия уменьшается, но трудоемкость возрастает.

2. Метод существенной выборки. При вычислении интегралов использовались случайные точки с плотностью $p(P)$. Желательно выбирать плотность $p(P)$ по возможности пропорциональной $|f(P)|$. Такой метод выбора $p(P)$ часто приводит к величинам $Z_0 = f(P)/p(P)$ с небольшими дисперсиями. В тех частях области G, в которых $|f(P)|$ больше и вклад в I_0 более существенен, будет выбираться больше случайных точек. И ещё один фактор — чем ближе $Z_0 = f(P)/p(P)$ к постоянной, тем меньше дисперсия величины $\mathbf{D}Z_0$.

3. Симметризация подынтегральной функции. Может применяться простая симметризация когда на интервале $[a, b]$ вместо функции $f(X)$ рассматривают симметризованную функцию

$$f_1(x) = 1/2\,[f(x) + f(a + b - x)],$$

интеграл которой по прежнему равен I_0.

В некоторых случаях интервал $[a, b]$ можно разбить на конечное число частей и на каждой из них использовать простую симметризацию. Такой метод называют сложной симметризацией.

К сожалению, различные методы симметризации, весьма наглядные и эффективные в одномерном случае, становятся громоздкими и трудно оцениваемыми при переходе к функциям многих переменных.

4. Двухэтапные схемы расчета. В некоторых задачах можно указать не одну «хорошую» оценку, а целое семейство, зависящее от параметров. Условием выбора наилучшего параметра обычно служит требование минимума дисперсии оценки (предполагается что время расчета слабо зависит от значений параметров).

Численный подход следующий: на первом этапе весьма грубо вычисляются дисперсии оценки при различных значениях параметров по небольшому количеству N испытаний. На втором этапе решается основная задача при помощи оценки с наилучшей системой параметров. Известные по первому этапу время счета и дисперсия позволяют довольно точно оценить объем работы, необходимый для достижения заданной вероятной ошибки.