

С.М. Ермаков

МЕТОД МОНТЕ-КАРЛО В
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКЕ

(ВВОДНЫЙ КУРС)

Санкт-Петербург
2009

УДК 519.21

ББК 517.8

Е 72

Рецензенты:

проф. В.Б. Невзоров (С.-Петербургский гос. ун-т),
проф. Ю.А. Сушков (С.-Петербургский гос. ун-т).

Книга посвящена быстро развивающемуся методу решения широкого круга прикладных задач — методу Монте-Карло. Автор известен своими исследованиями в этой области. Его монография "Метод Монте-Карло и смежные вопросы" выдержала 2 издания (1971, 1975 гг.) Двумя изданиями вышел также, написанный им совместно с Г.А.Михайловым, учебник "Статистическое моделирование". Настоящая книга может служить кратким и достаточно строгим введением в предмет. Вместе с тем она включает ряд новых результатов, относящихся к природе стохастических вычислительных методов, сравнению их с детерминированными аналогами и свойствам их параллелизма.

Книга адресована широкому кругу читателей, интересующимся современными проблемами математического моделирования и вычислительной математики.

© С.М. Ермаков, 2008

Оглавление

Предисловие	5
Глава 1. Моделирование равномерно распределенных случайных величин	7
1.1. Случайные числа (сведения из теории вероятностей)	7
1.2. Равномерно распределенные последовательности	13
1.3. Псевдослучайные числа	25
1.4. Библиографическое дополнение к главе 1	33
Глава 2. Моделирование случайных величин и процессов	35
2.1. Моделирование случайных величин	35
2.1.1. Формула обращения	36
2.1.2. Использование аппарата условных вероятностей	39
2.1.3. Метод Уокера	48
2.1.4. Битовые алгоритмы для моделирования дискретных распределений	49
2.2. Интегралы по траекториям	53
2.3. Библиографическое дополнение к главе 2	62
Глава 3. Вычисление интегралов	65
3.1. Метод Монте-Карло	65
3.1.1. Метод выделения главной части	67
3.1.2. Замена меры (метод существенной выборки)	68
3.1.3. Интегрирование по части переменных	73
3.1.4. Метод расслоенной выборки	74
3.1.5. Случайные интерполяционно-квадратурные формулы	76
3.2. Метод <i>квази</i> Монте-Карло	84
3.2.1. Некоторые сведения из теории кубатурных формул	84
3.2.2. Отклонение и его связь с нормой остатка	87
3.3. Библиографическое дополнение к главе 3	93
Глава 4. Вычисление суммы ряда Неймана	94
4.1. Несмещенные оценки	94
4.2. Вторые моменты, понижение дисперсии	102
4.3. Методы уменьшения дисперсии	105
4.3.1. Метод существенной выборки	105
4.3.2. Понижение порядка интегрирования	108
4.3.3. Ветвящиеся траектории	111

4.4. Замечания о трудоемкости. Квазислучайные числа	112
4.4.1. Сравнительная трудоемкость при решении систем линейных алгебраических уравнений	113
4.4.2. Метод <i>квази</i> Монте-Карло	116
4.5. Библиографическое дополнение к главе 4	117
Глава 5. Итерационные процессы. Стохастическая устойчивость	119
5.1. Разностные схемы и схема Неймана–Улама	119
5.2. Исследование стохастической устойчивости	125
5.3. Стохастическая устойчивость и параллелизм	133
5.3.1. Последовательные процедуры	133
5.3.2. Методы <i>квази</i> Монте-Карло	134
5.4. Обобщение на случай операторов	135
5.4.1. Схема Неймана–Улама	136
5.4.2. Рекуррентный алгоритм	138
5.5. Библиографическое дополнение к главе 5	142
Глава 6. Уравнения с полиномиальной нелинейностью	143
6.1. Ветвящиеся процессы и нелинейные уравнения	144
6.2. Уравнения с полиномиальной нелинейностью. Линеаризация .	153
6.3. Стохастическая устойчивость. Нелинейный случай	162
6.3.1. Стохастический аналог метода Ньютона	162
6.3.2. Квадратичная нелинейность. Метод искусственного хаоса	165
6.3.3. Искусственный хаос и метод Бёрда в газовой динамике	168
6.3.4. Заключительные замечания. Параллелизм	170
6.4. О решении разностного аналога уравнений Навье–Стокса . .	171
6.5. Библиографическое дополнение к главе 6	174
Заключение	175
Приложения	176
П.1. Простейшие оценки и условия согласования	176
П.2. Таблица простых чисел	177
П.3. Таблица числителей $r_j^{(l)}$	178
П.4. Примеры вычислений	182
Литература	188

Предисловие

В настоящее время метод Монте-Карло имеет необычайно широкий спектр приложений. В числе наиболее известных находятся

- задачи переноса излучений — ядерные реакторы, защита атмосферная оптика и др.;
- задачи газовой динамики — метод Бёрда, моделирование процессов коагуляции;
- задачи финансовой математики — моделирование управления ценными бумагами и рыночных ситуаций;
- задачи массового обслуживания — моделирование сложных производственных систем, систем связи и компьютерных сетей.

Этот список нетрудно расширить. По каждому из направлений имеются серьезные монографии. К их числу следует отнести, например, [12], [18], [22], [25]. Все направления быстро развиваются. Можно сказать, что прошло время, когда было легко составить более или менее полный обзор приложений метода Монте-Карло и не наступило время подводить итог развитию.

Данная книга предполагает осветить вопросы, общие для широкого круга приложений. Особое внимание уделяется стохастическим аналогам известных численных методов. Автор надеется, что читатель найдет в книге ответы на ряд принципиально важных вопросов, а именно:

- Как можно моделировать случайность детерминированными методами (на примере мультипликативного датчика псевдослучайных чисел)?
- В чем различие методов Монте-Карло и *квази*Монте-Карло?
- В каких задачах метод Монте-Карло полезен, а в каких его применять безусловно не следует (на примерах систем линейных алгебраических уравнений)?

Эти вопросы впервые достаточно полно освещаются в данной книге.

Книга может служить вводным курсом для широкого круга читателей, если:

- в главе 1 опустить доказательства большинства теорем (определения полезно усвоить!);
- главы 2, 3 и 4 читать полностью;

- главы 5 и 6 можно опустить — они носят специальный характер. В особенности это относится к заключительным параграфам этих глав, которые выходят за рамки вводного курса.

Автор стремился везде, где это возможно, использовать элементарные средства для доказательства основных теорем. Что касается отбора материала, то рассматривались лишь наиболее употребительные методы, позволяющие решить как можно более широкий круг задач. Этим обусловлено предпочтение, отданное дискретным задачам. Непрерывный случай в более сложном изложении читатель найдет, например, в книгах Ермакова, Некруткина, Сипина [45]; Сабельфельда [65]; Михайлова [57].

В данной книге также не рассматриваются вопросы, связанные с решением стохастических дифференциальных уравнений. Эта проблематика требует отдельной монографии.

Новые результаты, вошедшие в книгу, обсуждались на Кафедре статистического моделирования Математико-механического факультета Санкт-Петербургского государственного университета. Автор благодарен своим коллегам за ряд полезных замечаний и помощь в подготовке текста. Особую признательность автор выражает С. Н. Леора. Без ее участия книга еще долго не вышла бы из печати. Наконец, нужно отметить решающую роль Российского фонда фундаментальных исследований (гранты № 08-01-00194а, № 05-01-00865а), оказавшего поддержку в работе автору и его коллегам.

Глава 1

Моделирование равномерно распределенных случайных величин

Моделирование случайности с использованием компьютера имеет обширную сферу приложений. Почти каждая игра, которую мы проигрываем на компьютере, использует элемент случайности. С помощью так называемого датчика случайных чисел имитируется движение молекул в сложных средах, изучаются производственные процессы, экономическая и финансовая ситуация и многое другое. Метод Монте-Карло — это метод, использующий случайные числа для решения разнообразных задач. Случайные числа можно получать с помощью рулетки, что собственно и происходит в игорных заведениях Монте-Карло. Так возникло название метода, а его развитие связано в первую очередь с расчетами атомной бомбы и ядерных реакторов. Проблемы в этой области до сих пор весьма ощутимо влияют на научные исследования, посвященные методу Монте-Карло.

Имеется обширная монографическая и учебная литература на русском языке, где, в частности, освещена предыстория метода. Мы не будем далее останавливаться на этих вопросах, а рассмотрим математические аспекты имитации случайности.

1.1. Случайные числа (сведения из теории вероятностей)

Напомним некоторые простейшие сведения из теории вероятностей, относящиеся в основном к равномерно распределенным случайным величинам.

Простейшие случайные величины или случайные числа

Простейшими будем считать дискретные или непрерывные равномерно распределенные случайные величины: величину β , такую что она

принимает одно из r значений $0, 1, \dots, r-1$ с равными вероятностями

$$P\{\beta = k; k = 0, \dots, r-1\} = 1/r, \quad (1.1.1)$$

а также величину α , равномерно распределенную на промежутке $[0, 1]$, то есть такую, что она может принимать любое вещественное значение из $[0, 1]$, и вероятность того, что она примет значение из некоторого интервала $\Delta \in [0, 1]$, равна длине $|\Delta|$ этого интервала. Плотность $\phi(x)$ на $[0, 1]$ случайной величины α выражается равенством

$$\phi(x) = \begin{cases} 1, & x \in [0, 1], \\ 0, & x \notin [0, 1]. \end{cases} \quad (1.1.2)$$

Для дискретной величины особый интерес представляет $r = 2$. В этом случае β называют случайным битом.

Независимые реализации случайной величины β называют также (особенно при $r = 10$) случайными цифрами, а числа, составленные из этих случайных цифр, — случайными числами.

Некоторые элементарные факты, относящиеся к случайным числам, будут важны в дальнейшем изложении. Мы сформулируем их в виде свойств.

Случайные величины α и β связаны между собой. Имеет место следующее утверждение.

Свойство 1. Если $0_r \beta_1 \beta_2 \dots \beta_n \dots$ — бесконечная r -ичная дробь *) и β_i — независимые в совокупности случайные величины с распределением (1.1.1), то случайная величина $\alpha = 0_r \beta_1 \dots \beta_n \dots$ равномерно распределена на промежутке $[0, 1]$. Обратно: если α равномерно распределена на $[0, 1]$ и $0_r \beta_1 \dots \beta_n \dots$ — ее представление в виде r -ичной дроби, то $\beta_1, \dots, \beta_n, \dots$ независимы в совокупности и имеют распределение (1.1.1). $\square$

Отметим, что аналогичные утверждения могут быть доказаны в случае представления дроби в смешанной системе счисления.

Дальнейшее будет касаться случайных векторов, компоненты которых независимы и равномерно распределены.

Свойство 2. Пусть $\alpha_1, \alpha_2, \dots, \alpha_n, \dots$ — последовательность равномерно распределенных независимых реализаций случайной величины α (с плотностью (1.1.2)), тогда случайные векторы

$$\Xi_i = (\alpha_{(i-1)s+1}, \dots, \alpha_{is}), \quad i = 1, 2, \dots,$$

*) То есть $\sum_{k=1}^{\infty} \beta_k / r^k$ — здесь r играет роль запятой в десятичном представлении ($0, \dots = 0_{10}$).

независимы и равномерно распределены в единичном s -мерном гиперкубе $D_s = [0, 1]^s$. $\square$

Свойство 2а. При тех же условиях случайные векторы

$$(\alpha_i, \alpha_{i+1}, \dots, \alpha_{i+s-1}), \quad i = 1, 2, \dots, \quad (1.1.3)$$

равномерно распределены в единичном s -мерном гиперкубе. $\square$

Свойство 3. Если мы произвольным (но не зависящим от значений $\alpha_1, \alpha_2, \dots$) образом изменим порядок нумерации элементов последовательности $\|\alpha_i\|_1^\infty$, то свойство 2 также будет иметь место $^*)$. $\square$

Далее сформулируем утверждения, являющиеся частными случаями закона больших чисел и центральной предельной теоремы соответственно, рассматривая их как свойства равномерно распределенных случайных величин. Пусть Ξ_i — векторы с равномерно распределенными и независимыми компонентами (фигурирующими в свойствах 2 и 3).

Тогда имеет место

Свойство 4. (Частный случай усиленного закона больших чисел.)

Если $f(X)$ — любая интегрируемая (по Лебегу) в единичном гиперкубе $D_s = [0, 1]^s$ функция, $X = (x_1, \dots, x_s)$, то с вероятностью 1

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(\Xi_i) = \int_{D_s} f(X) dX \quad (1.1.4)$$

для любого $s = 1, 2, \dots$ $\square$

Свойство 5. (Частный случай центральной предельной теоремы.)

Для любой интегрируемой с квадратом в D_s функции $f(X)$ при любом натуральном s имеем

$$\lim_{N \rightarrow \infty} \mathbf{P} \left(\frac{\sqrt{N}}{\sigma_f} \left| \frac{1}{N} \sum_{i=1}^N f(\Xi_i) - \int_{D_s} f(X) dX \right| < y \right) = \sqrt{\frac{2}{\pi}} \int_0^y e^{-u^2} du,$$

где

$$\sigma_f^2 = \int_{D_s} f^2(X) dX - \left(\int_{D_s} f(X) dX \right)^2. \quad (1.1.5)$$

$\square$

$^*)$ Строго говоря, речь идет об изменении нумерации в любом конечном подмножестве элементов последовательности $\alpha_1, \alpha_2, \dots$

Два последних свойства указывают на возможность практического использования случайных чисел (если мы умеем их получать) в вычислительной математике. Действительно, интеграл

$$J = \int_{D_s} f(X) dX$$

при достаточно большом N может быть приближен (с вероятностью 1) средним $S_N[f] = \frac{1}{N} \sum_{i=1}^N f(\Xi_i)$. Если же интеграл $\int_{D_s} f^2(X) dX$ конечен, то с вероятностью p и также достаточно большим N имеет место неравенство

$$|J - S_N[f]| \leq \frac{\gamma_p \sigma_f}{\sqrt{N}}, \quad (1.1.6)$$

γ_p определяется соотношением

$$\sqrt{\frac{2}{\pi}} \int_0^{\gamma_p} e^{-u^2/2} du = p.$$

В действительности, располагая случайными числами, мы можем использовать предельные законы теории вероятностей в гораздо более общей форме. Например, можно получить предельное распределение среднего $S_N[f]$, когда существует интеграл

$$\int_{D_s} |f(X)|^r dX, \quad 1 < r < 2,$$

и оценить погрешность $|J - S_N[f]|$. Она убывает в этом случае медленнее чем $O(N^{-1/2})$. Кроме того, можно рассмотреть задачу о вычислении вероятностным способом интеграла по вероятностной мере μ . Этот вопрос будет обсуждаться в следующих главах.

Однако вернемся к вопросу о случайных числах. Попытаемся получить их, подбрасывая монету (если она идеальна в том смысле, что орел (1) и решка (0) “выпадают” с равными вероятностями) или с помощью рулетки (отсюда, собственно, произошло название “метод Монте-Карло”), также идеальной.

Один из подходов, позволяющих использовать некую последовательность чисел (r -ичных цифр) в приложениях, — применение аппарата статистических тестов (проверки гипотез).

Имея заданную последовательность чисел $\alpha_1, \alpha_2, \dots$, мы можем высказывать гипотезы о том, что они случайны, независимы и имеют заданное распределение. Аппарат проверки статистических гипотез позволяет проверить, не противоречат ли наши предположения наблюдаемым значениям чисел. Если такое противоречие имеет место

со значительной вероятностью, то соответствующая гипотеза отвергается.

Удовлетворение данной последовательности тестам является необходимым условием их использования в приложениях.

В качестве примеров критериев согласия для проверки статистических гипотез приведем критерии χ^2 и Колмогорова.

Пусть нами высказана гипотеза H_0 : случайная величина α имеет равномерное распределение на промежутке $[0, 1]$. Критерий χ^2 строится следующим образом. Выбирается m взаимно исключающих событий $A_1, \dots, A_m$ так, что появление каждой реализации α_i случайной величины α означает наступление одного из них. Подсчитывается число n_k наступления k -го события, $\sum_{k=1}^m n_k = N$. Если $P(A_k)$ — вероятность события, вычисленная в предположении, что H_0 верна, то вычисляем значение критерия χ^2 :

$$\chi_{m-1}^2 = \sum_{k=1}^m C_k \left(\frac{n_k}{N} - P(A_k) \right)^2, \quad C_k > 0. \quad (1.1.7)$$

В нашем случае отрезок $[0, 1]$ разбивается на непересекающиеся отрезки Δ_k , и для каждого из них подсчитывается число событий $A_k: \alpha_i \in \Delta_k, \quad i = 1, \dots, N$. Теоретическая вероятность события A_k есть длина отрезка Δ_k , $P(A_k) = |\Delta_k|$. Обычно выбирают $|\Delta| = 1/m$.

Критерий (1.1.7) можно также вычислить для последовательности цифр β_i (в r -ичной системе счисления). В предположении, что каждая цифра равновероятна, обычно полагают

$$A_k = \{\beta_i = k\}, \quad k = 0, 1, \dots, r-1, P(A_k) = 1/r, \quad m = r.$$

Справедлива следующая теорема.

Теорема 1.1. При выборе констант $C_k = N/P(A_k)$ величина χ_{m-1}^2 имеет при $N \rightarrow \infty$ распределение χ^2 с $m-1$ степенью свободы. $\square$

Для распределения χ_m^2 существуют подробные таблицы.

Важно отметить, что есть очень много возможностей выбора событий A_k в предположении справедливости H_0 . Можно проверять события, состоящие в том, что появление всевозможных пар β_{2j-1}, β_{2j} равновероятно. То же относительно троек и т. д. Аналогично мы можем проверить равномерность распределения пар, троек и др. в непрерывном случае. При этом легко видеть, что равновероятность появления каждого из значений $\beta_k = k$ в ряду $\beta_1, \dots, \beta_N$ не гарантирует равновероятность появления всевозможных пар. (Сравните свойство 2 равномерно распределенных случайных величин.)

Критерий Колмогорова

Если $\alpha_1, \alpha_2, \dots, \alpha_N$ — случайные точки на $\mathbf{R}^1$, то функцию

$$\hat{F}_N(x) = \frac{1}{N} \sum_{k=1}^N \chi_{(-\infty, x]}(\alpha_k) = \frac{1}{N} \sum_{k=1}^N \mathbf{1}(x_k \leq x), \quad k = 1, \dots, N, \quad (1.1.8)$$

где $\chi_{(a,b]}(x)$ — характеристическая функция отрезка $(a, b]$, называют эмпирической функцией распределения. Критерием Колмогорова называют случайную величину

$$D_N = \sup_x |\hat{F}_N(x) - F(x)|, \quad (1.1.9)$$

где $F(x)$ — теоретическая функция распределения α . В интересующем нас случае $F(x) = x$. Известно предельное при $N \rightarrow \infty$ распределение D_N , которое не зависит от $F(x)$, что дает возможность проверять гипотезу о распределении случайных величин. В нашем случае это гипотеза о равномерности распределения.

Для случайного вектора в s -мерном гиперкубе можно рассматривать эмпирическую функцию распределения каждой компоненты (в предположении, что гипотеза H_0 верна: $F(x) = x$ — это одна и та же функция распределения для каждой из компонент), а также эмпирические функции их условного распределения, соответствующие условным функциям распределения, при $s = 2$

$$F(x, y) = P(\alpha \leq x/\alpha' \leq y). \quad (1.1.10)$$

На их базе могут строиться критерии согласия для проверки гипотезы H_0 .

В литературе неоднократно описывались системы тестов, которые авторы считают достаточными для исследования последовательностей чисел, используемых в качестве случайных и равномерно распределенных (например, [53], [17]). Не останавливаясь подробнее на этом вопросе, отметим, что проверки с помощью статистических тестов полезны и необходимы, но исходный выбор алгоритма, приближенно имитирующего случайность, нуждается в непосредственном теоретическом обосновании. Далее мы рассмотрим один из подходов к такого рода обоснованию в рамках теоретико-числовых методов.

Псевдослучайные числа

Как уже отмечалось, многие приложения предполагают, что независимые реализации случайной величины с равномерным распределением используются компьютерными программами.

Удобным методом получения таких реализаций, очевидно, является сравнительно простой алгоритм. Однако здесь мы имеем дело с явным противоречием: детерминированный алгоритм, строго говоря, не может быть источником случайных чисел. Можно предположить, что алгоритм, вычисляющий числа, “похожие” на случайные, должен быть очень сложным. Действительно, исследования Н.А. Колмогорова и его последователей показали, что алгоритм, дающий числа, которые не могут быть отличимы от случайных (с помощью каких-либо тестов), должен иметь бесконечную сложность (см., например, [78]). По этой причине следует ставить задачу о конструкции алгоритма, вычисляющего числа, обладающие лишь некоторыми свойствами случайных чисел или обладающие этими свойствами приближенно (в некотором строго определенном смысле).

Определение 1.1. Назовем последовательность чисел $\alpha_1, \dots, \alpha_N, \dots$ псевдослучайной, если эти числа вычисляются с помощью детерминированного алгоритма (программы) и приближенно обладают свойствами 1–5 случайных чисел.

Решение проблемы построения такого алгоритма, называемого датчиком случайных чисел, лежит вне рамок теории вероятностей. Здесь необходимо привлечение средств теории чисел. Ниже излагаются некоторые нужные нам факты из теории равномерно распределенных в теоретико-числовом смысле последовательностей.

1.2. Равномерно распределенные последовательности вещественных чисел

Необходимо четко уяснить себе, что существует принципиальное различие между теоретико-вероятностным и теоретико-числовым подходами. В рамках теории чисел мы забываем временно о случайности. Мы имеем дело с последовательностями (может быть, бесконечными) целых или вещественных чисел.

Пусть $\mathcal{E} = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, \dots$ — последовательность вещественных чисел из промежутка $[0, 1]$. В рамках теории чисел можно рассматривать аналоги свойств 1–5, сформулированных ранее для последовательности независимых реализаций равномерно распределенной случайной величины.

Так, например, для интегрируемой на $[0, 1]$ функции $f(x)$ мы можем рассматривать предел

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(\varepsilon_k). \quad (1.2.1)$$

Если для любой интегрируемой f этот предел будет равен интегралу $\int_0^1 f(x) dx$, то мы вправе говорить об аналоге закона больших чисел (свойство 4). Однако среди всех интегрируемых по Лебегу функций можно указать такие, что они просто не определены в точках $\varepsilon_1, \varepsilon_2, \dots$ последовательности $\mathcal{E}$, множество которых имеет нулевую меру. Для этих функций равенство

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(\varepsilon_k) = \int_0^1 f(x) dx \quad (1.2.2)$$

не обязано выполняться.

Известно, однако, что существуют последовательности, такие что для любой интегрируемой по Риману функции f равенство (1.2.2) выполняется.

Определение 1.2. Последовательность $(\varepsilon_i)_{i=1}^\infty$ вещественных чисел, принадлежащих отрезку $[0, 1]$, называется равномерно распределенной (в теоретико-числовом смысле) на этом отрезке, если для любой интегрируемой по Риману функции $f(x)$ имеет место равенство (1.2.2).

Поскольку ступенчатые функции и тригонометрические функции (теорема Вейерштрасса) всюду плотны в множестве интегрируемых по Риману функций, то определение 1.2 равносильно следующим определениям (соответствующие доказательства имеются в доступной литературе).

Определение 1.2а. Последовательность $(\varepsilon_i)_{i=1}^\infty$ называется равномерно распределенной, если для каждого подынтервала $\Delta \subset [0, 1]$ выполняется равенство

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \chi_\Delta(\varepsilon_i) = |\Delta|,$$

где

$$\chi_\Delta(x) = \begin{cases} 1, & x \in \Delta, \\ 0, & x \notin \Delta. \end{cases}$$

Определение 1.26. (Критерий Вейля.) Последовательность $(\varepsilon_l)_1^\infty$ равномерно распределена, если для каждого целого, отличного от нуля k выполняется равенство

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=1}^N e^{2\pi i k \varepsilon_l} = 0.$$

Справедлива следующая теорема.

Теорема 1.2. Определения 1.2, 1.2а и 1.2б равносильны [52]. $\square$

Из этой теоремы (определение 1.2б) следует

Теорема 1.3. Равномерно распределенные последовательности существуют.

Доказательство. В качестве доказательства мы укажем конкретный пример такой последовательности (также следуя Вейлю). Пусть

$$\varepsilon_n = \{n\beta\}, \quad n = 1, 2, \dots, \quad (1.2.3)$$

где β — иррациональное число, а фигурные скобки обозначают дробную долю заключенного в них числа a . Последовательность ε_n равномерно распределена. Проверим критерий Вейля:

$$S_N(\mathcal{E}) = \sum_{n=1}^N e^{2\pi i k \{n\beta\}} = \sum_{n=1}^N e^{2\pi i k n \beta} = e^{2\pi i k \beta} \cdot \frac{1 - e^{2\pi i k N \beta}}{1 - e^{2\pi i k \beta}}$$

и

$$|S_N(\mathcal{E})| \leq \left| e^{2\pi i k \beta} \cdot \frac{1 + |e^{2\pi i k N \beta}|}{1 - e^{2\pi i k \beta}} \right|,$$

что после элементарных выкладок дает

$$|S_N(\mathcal{E})| \leq \frac{1}{|\sin \pi k \beta|}.$$

Предположение об иррациональности β обеспечивало во всех случаях отличие от нуля знаменателя в правой части неравенств.

Таким образом, имеем

$$\frac{1}{N} |S_N(\mathcal{E})| \leq \frac{1}{N |\sin \pi k \beta|}$$

и при любом целом $k \neq 0$

$$\lim_{N \rightarrow \infty} N^{-1} |S_N(\mathcal{E})| = 0,$$

что и доказывает равномерную распределенность последовательности $\mathcal{E}$. $\square$

В действительности, равномерная распределенность — довольно слабое свойство, и равномерно распределенных последовательностей “очень много”. Мы приведем далее так называемую метрическую теорему (см. [52]), относящуюся к почти всем значениям величины x из $[0, 1]$.

Теорема 1.4. Пусть $a_1, a_2, \dots, a_n, \dots$ — последовательность различных целых чисел, тогда последовательность

$$\varepsilon_n = \{a_n x\} \quad (1.2.4)$$

равномерно распределена для почти всех x из промежутка $[0, 1]$.

Доказательство. Зафиксируем целое $k \neq 0$ и положим

$$S(N, x) = \frac{1}{N} \sum_{n=1}^N e^{2\pi i k a_n x}, \quad N \geq 1, \quad x \in [0, 1].$$

Домножая $S(N, x)$ на $\overline{S(N, x)}$ — комплексно сопряженную величину, имеем

$$S(N, x) \overline{S(N, x)} = |S(N, x)|^2 = \frac{1}{N^2} \sum_{m, n=1}^N e^{2\pi i k (a_m - a_n) x}.$$

Интегрируя это выражение по x от 0 до 1 и замечая, что отличными от нуля (равными единице) будут лишь слагаемые при $m = n$, получим

$$\int_0^1 |S(N, x)|^2 dx = \frac{1}{N}.$$

Следовательно,

$$\sum_{N=1}^{\infty} \int_0^1 |S(N^2, x)|^2 dx = \sum_{N=1}^{\infty} \frac{1}{N^2} < \infty.$$

По лемме Фату получим

$$\int_0^1 \sum_{N=1}^{\infty} |S(N^2, x)|^2 dx < \infty.$$

Отсюда следует, что ряд под знаком интеграла должен сходиться для почти всех x и его общий член для почти всех x должен стремиться к нулю:

$$\lim_{N \rightarrow \infty} S(N^2, x) = 0, \quad x \in [0, 1].$$

Для целого положительного N можно указать такое целое положительное m , что $m^2 \leq N < (m+1)^2$. Оценивая грубо $|S(N, x)|$, получим

$$\begin{aligned} |S(N, x)| &\leq |S(m^2, x)| + \frac{1}{N} \left| \sum_{n=m^2+1}^N e^{2\pi i k a_n x} \right| \leq |S(m^2, x)| + \frac{2m}{N} \leq \\ &\leq |S(m^2, x)| + \frac{2}{\sqrt{N}}, \quad \text{поскольку } N \leq m^2 + 2m. \end{aligned}$$

Отсюда $\lim_{N \rightarrow \infty} S(N, x) = 0$ для почти всех x из $[0, 1]$.

Исключительное множество зависит от k , но сумма счетного множества множеств нулевой меры имеет меру 0. Это и завершает доказательство теоремы. $\square$

Аналогично одномерному случаю вводится понятие равномерно распределенной последовательности векторов.

Определение 1.3. Последовательность s -мерных векторов

$$\mathbf{e}_i = \|(e_i^{(1)}, \dots, e_i^{(s)})\|_{i=1}^\infty, \quad 0 \leq e_i^{(j)} \leq 1, \quad j = 1, \dots, s,$$

называется равномерно распределенной в $D_s = [0, 1]^s$ — единичном s -мерном гиперкубе, если для любой интегрируемой по Риману в D_s функции $f(X)$, $X = (x_1, \dots, x_s)$, выполняется предельное равенство

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\mathbf{e}_n) = \int_{D_s} f(X) dX.$$

Имеет место многомерный аналог критерия Вейля.

Теорема 1.4а. Для того чтобы последовательность $\mathbf{e}_i$ была равномерно распределена в D_s , необходимо и достаточно, чтобы для любого набора s целых чисел $k_1, \dots, k_s$, не равных одновременно нулю, выполнялось равенство

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i (k_1 e_n^{(1)} + \dots + k_s e_n^{(s)})} = 0. \quad (1.2.5)$$

$\square$

Если любой s -мерный прямоугольный параллелепипед, лежащий в D_s , обозначить Δ_s , то будет иметь место очевидный аналог теоремы 1.2.

Если $\beta_1, \dots, \beta_s$ — некоторые независимые иррациональные числа, то есть такие, что никакая их линейная комбинация с целыми коэффициентами не является рациональным числом, то легко видеть, что для последовательности векторов

$$\mathbf{e}_n = (\{n\beta_1\}, \{n\beta_2\}, \dots, \{n\beta_s\})$$

выполнен критерий Вейля (1.2.5), то есть она равномерно распределена в D_s .

Продолжая аналогию между равномерно распределенными случайными величинами и равномерно распределенными последовательностями, замечаем, что желательным для нас является выполнение аналогов свойств 2 и 3. В теории чисел фигурирует аналог свойства 2а.

Определение 1.4. Последовательность $\|\varepsilon_i\|$ называется s -равномерно распределенной, если последовательность векторов $(\varepsilon_i, \varepsilon_{i+1}, \dots, \varepsilon_{i+s-1})$, $i = 1, 2, \dots$, равномерно распределена в D_s .

Непосредственно с помощью критерия Вейля легко установить следующее.

В случае если последовательность $\|\varepsilon_i\|$ s -равномерно распределена и $r \leq s$, то она и r -равномерно распределена. Обратное, вообще говоря, неверно.

О дальнейших свойствах s -равномерно распределенных в смысле этого определения последовательностей см. [24], [52]. Мы дадим другое определение s -равномерно распределенных последовательностей.

Пусть $\varepsilon = \|\varepsilon_i\|_1^\infty$ — последовательность вещественных чисел из $[0, 1]$. Образует векторы размерности s :

$$\mathbf{e}_k^{(s)} = (\varepsilon_{(k-1)s+1}, \dots, \varepsilon_{ks}), \quad k = 1, 2, \dots$$

Определение 1.5. Назовем последовательность $\|\varepsilon_i\|$ — s -независимо равномерно (s -н-равномерно) распределенной, если векторы $\mathbf{e}_k^{(s)}$ равномерно распределены в D_s .

Такое определение менее удобно для исследований в области теории чисел, но более естественно в теории метода Монте-Карло.

Неизвестно, справедливо ли для любых s -н-равномерно распределенных последовательностей утверждение, что они также являются

$(s - 1)$ -н-равномерно распределенными. (Это утверждение очевидно лишь для $s = 2$.)

Однако следующая теорема имеет место.

Теорема 1.5. Из s -н-равномерной распределенности последовательности не следует, вообще говоря, ее $(s + 1)$ -н-равномерная распределенность.

Доказательство. Для доказательства данного утверждения достаточно рассмотреть простой пример.

Пусть $s = 1$ и $\varepsilon_k = \{k\beta\}$, где β иррационально. Не умаляя общности, считаем, что $\beta < 1$. Построим последовательность $\mathbf{e}_l^{(2)}$

$$\mathbf{e}_l^{(2)} = (\{(2l - 1)\beta\}, \{2l\beta\}), \quad l = 1, 2, \dots \quad (1.2.6)$$

Покажем, что векторы $\mathbf{e}_l^{(2)}$ не распределены равномерно в квадрате. Действительно, в случае их равномерной распределенности должно, в частности, выполняться равенство

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=1}^N \chi_{\Delta}(\mathbf{e}_l^{(2)}) = \int_{\Delta} dx dy = \frac{1}{2}, \quad (1.2.7)$$

где $\Delta = \{x, y : 0 \leq x \leq y \leq 1\}$, а χ_{Δ} — характеристическая функция этого треугольника. Равенство (1.2.7) означает также, что неравенство

$$\{(2l - 1)\beta\} \leq \{2l\beta\} \quad (1.2.8)$$

должно выполняться в пределе для половины членов последовательности ε_k . Это неравенство является условием нахождения точки $\mathbf{e}_l^{(2)}$ в треугольнике. Однако

$$\{2l\beta\} = \{\{(2l - 1)\beta\} + \beta\},$$

где

$$\{\{(2l - 1)\beta\} + \beta\} = \begin{cases} (2l - 1)\beta + \beta, & \text{если } \{(2l - 1)\beta\} + \beta < 1, \\ (2l - 1)\beta + \beta - 1, & \text{если } \{(2l - 1)\beta\} + \beta > 1. \end{cases}$$

При этом второй случай не может быть реализован, так как из (1.2.8) следует

$$\{(2l - 1)\beta\} \leq \{2l\beta\} = \{(2l - 1)\beta\} + \beta - 1, \quad \text{то есть } \beta - 1 > 0 \text{ и } \beta > 1,$$

что противоречит предположению $\beta < 1$. Таким образом, неравенство (1.2.8) имеет место, только если $\{(2l-1)\beta\} < 1 - \beta$. В силу равномерной распределенности ^{*)}

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=1}^N \chi_{\Delta}(\mathbf{e}_l^{(2)}) = (1 - \beta) \neq \frac{1}{2},$$

β — иррационально.

Отсюда и следует наше утверждение. $\square$

В теории чисел показано, что s -равномерно распределенные (в смысле определения 1.4) последовательности $\rho_k^{(s)}$, $k = 1, 2, \dots$, могут быть получены с помощью формулы

$$\rho_k^{(s)} = \{k^s + P_{s-1}(k)\beta\},$$

где P_{s-1} — многочлен степени $s-1$ с целыми коэффициентами [10].

Последовательности, s -равномерно распределенные при любом s , носят название вполне равномерно распределенных. Очевидно, свойство вполне равномерной распределенности является важным. Им обладает (с вероятностью 1) последовательность реализаций равномерно распределенных случайных величин. Можно ожидать, что, заменив в выражении для $\rho_k^{(1)}$ многочлен на экспоненту, мы получим последовательность со свойствами, близкими к характеристикам вполне равномерной распределенности. Обсудим далее некоторые свойства таких последовательностей.

Некоторые свойства дробных долей показательной функции

Будем рассматривать последовательность вещественных чисел вида

$$\varepsilon_k = \{M^k \varepsilon_0\} \quad \text{или} \quad \varepsilon_k = \{M \varepsilon_{k-1}\}. \quad (1.2.9)$$

Для дальнейшего нам понадобится теорема (см. [16], [35], [70]), которую мы приведем в более простой форме, чем это сделано в [35].

Теорема 1.6. Пусть $f(X)$ — функция, определенная и интегрируемая по Риману в единичном s -мерном гиперкубе D_s .

Положим

$$\phi_M(x) = f(x, \{M_1 x\}, \dots, \{M_1 \dots M_{s-1} x\}),$$

^{*)} С помощью критерия Вейля легко проверить, что последовательности $\{2l\beta\}$ и $\{(2l-1)\beta\}$ равномерно распределены на отрезке $(0, 1)$, если β иррационально.

где $M = \min(M_1, \dots, M_s)$ и $M_1, \dots, M_{s-1}$ — натуральные числа больше единицы. Справедливо равенство

$$\lim_{M \rightarrow \infty} \int_0^1 \phi_M(x) dx = \int_{D_s} f(X) dX. \quad (1.2.10)$$

Доказательство. Разобьем отрезок $(0, 1)$ на M_1 равных интервалов, положим $M_1 x = j_1 + y_1$ на интервале с номером j_1 . Имеем

$$\begin{aligned} \int_0^1 \phi_M(x) dx &= \sum_{j_1=0}^{M_1-1} \int_{j_1/M_1}^{(j_1+1)/M_1} \phi_M(x) dx = \\ &= \frac{1}{M_1} \sum_{j_1=0}^{M_1-1} \int_0^1 f\left(\frac{j_1 + y_1}{M_1}, y_1, \dots, \{M_2 \dots M_{s-1} y_1\}\right) dy_1. \end{aligned}$$

Аналогично разбивая каждый из полученных интегралов на сумму M_2 слагаемых и в каждом из них делая замену переменных

$$M_2 y_1 = j_2 + y_2,$$

получим

$$\begin{aligned} &\int_0^1 \phi_M(x) dx = \\ &= \frac{1}{M_1 M_2} \sum_{j_1=0}^{M_1-1} \sum_{j_2=0}^{M_2-1} \int_0^1 f\left(\frac{j_1 + y_1}{M_1}, \frac{j_2 + y_2}{M_2}, y_2, \dots, \{M_3 \dots M_{s-1} y_2\}\right) dy_2. \end{aligned}$$

И после повторения этого действия s раз (в последний раз отрезок разбиваем на M частей) имеем

$$\begin{aligned} \int_0^1 \phi_M(x) dx &= \frac{1}{M M_1 \dots M_{s-1}} \times \\ &\times \sum_{j_1=0}^{M_1-1} \dots \sum_{j_s=0}^{M-1} \int_0^1 f\left(\frac{j_1 + y_1}{M_1}, \dots, \frac{j_{s-1} + y_{s-1}}{M_{s-1}}, \frac{j_s + y_s}{M}\right) dy_s. \end{aligned}$$

Обозначим J_M правую часть последнего равенства, а $\overline{m}(j_1, \dots, j_s)$ и $\underline{m}(j_1, \dots, j_s)$ — соответственно верхнюю и нижнюю грани функции $f(X)$ в параллелепипеде

$$j_k/M_k \leq x_k \leq (j_k + 1)/M_k, \quad k = 1, \dots, s; \quad M_s = M,$$

тогда очевидным образом получим

$$\begin{aligned} \frac{1}{M_1 \dots M_{s-1} M} \sum_{j_1=0}^{M_1-1} \dots \sum_{j_s=0}^{M_s-1} \underline{m}(j_1, \dots, j_s) &< J_M \leq \\ &\leq \frac{1}{M_1 \dots M_{s-1} M_s} \sum_{j_1=0}^{M_1} \dots \sum_{j_s=0}^{M_s} \overline{m}(j_1, \dots, j_s). \end{aligned}$$

В правой и левой частях этого неравенства находятся верхняя $\overline{S}(f, M)$ и нижняя $\underline{S}(f, M)$ суммы Дарбу, которые в силу предположенной интегрируемости по Риману функции f имеют при $M \rightarrow \infty$ своим общим пределом $\int_{D_s} f(X) dX$, то есть

$$\int_0^1 \phi_M(x) dx \text{ при } M \rightarrow \infty \text{ есть } \int_{D_s} f(X) dX. \quad \square$$

Замечание. Результат теоремы тривиальным образом не зависит от порядка нумерации независимых переменных. $\square$

Теперь мы можем убедиться в том, что для последовательности дробных долей показательной функции асимптотически при $M \rightarrow \infty$ выполняются многие свойства последовательности независимых реализаций случайной величины, равномерно распределенной на промежутке $[0, 1]$.

Теорема 1.7. Пусть $f(X)$ — произвольная интегрируемая по Риману в D_s функция, $\alpha_l = \{M^{l-1}\beta\}$, $\beta \in [0, 1]$ и $X_k = (\alpha_{(k-1)s+1}, \dots, \alpha_{ks})$. Тогда для почти всех β из $(0, 1)$ справедливо равенство

$$\lim_{M \rightarrow \infty} \left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(X_k) \right) = \int_{D_s} f(X) dX. \quad (1.2.11)$$

Доказательство. Действительно, пусть M фиксировано. Имеем

$$f(X_k) = f(\{M^{(k-1)s+1}\beta\}, \dots, \{M^{ks}\beta\}) = \phi_M(\{M^{(k-1)s}\beta\}).$$

Из теоремы 1.4 следует, что для почти всех β из $[0, 1)$ последовательность $Y_k = \{M^{(k-1)s}\beta\}$ распределена равномерно и мы имеем

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(X_k) = \int_0^1 \phi_M(x) dx. \quad (1.2.12)$$

Применяя к (1.2.12) теорему Сю [16], получаем требуемое. $\square$

В теореме 1.7 векторы X_k были образованы из элементов последовательности $\alpha_1, \dots, \alpha_t$, $\alpha_t = \{M^t x\}$, взятых в определенном фиксированном порядке. Пусть теперь $l_1^k, \dots, l_s^k$ — некоторые различные неотрицательные целые числа. Рассмотрим векторы $Y_k = (\alpha_{l_1^k}, \dots, \alpha_{l_s^k})$, $k = 1, 2, \dots, N$. Имеет место следующий результат.

Теорема 1.7а. Для последовательности векторов Y_k асимптотически по M выполняется критерий Вейля для почти всех x из $[0, 1]$.

Доказательство. Во многом аналогично доказательству теоремы 1.4. Зафиксируем набор целых чисел $k_1, \dots, k_s$, не равных одновременно нулю, и составим сумму

$$S(N, x) = \frac{1}{N} \sum_{n=1}^N \exp [2\pi i (k_1 \alpha_{l_1^n} + k_2 \alpha_{l_2^n} + \dots + k_s \alpha_{l_s^n})].$$

Умножая $S(N, x)$ на сопряженную к ней, имеем

$$|S(N, x)|^2 = \frac{1}{N^2} \sum_{m, n=1}^N \exp [2\pi i (k_1 (M^{l_1^n} - M^{l_1^m}) + \dots + k_s (M^{l_s^n} - M^{l_s^m})) x].$$

Выделяя члены суммы, соответствующие $m = n$, получим

$$|S(N, x)|^2 = \frac{1}{N} + \frac{1}{N^2} \sum_{\substack{m \neq n \\ m, n=1}}^N \exp [2\pi i x (k_1 (M^{l_1^n} - M^{l_1^m}) + \dots + k_s (M^{l_s^n} - M^{l_s^m}))].$$

Проинтегрируем полученное выражение по x и перейдем к пределу:

$$\begin{aligned} \lim_{M \rightarrow \infty} \int_0^1 |S(N, x)|^2 dx &= \frac{1}{N} + \frac{1}{N^2} \times \\ &\times \sum_{\substack{m \neq n \\ m, n=1}}^N \lim_{M \rightarrow \infty} \int_0^1 [2\pi i x (k_1 (M^{l_1^n} - M^{l_1^m}) + \dots + k_s (M^{l_s^n} - M^{l_s^m}))]. \end{aligned} \quad (1.2.13)$$

Заметим, что в каждом слагаемом суммы правой части (1.2.13) подынтегральное выражение есть некая $\varphi_M(x)$ (см. теорему Сю) для функции $2s$ переменных $\exp [2\pi i (k_1 (x_1 - y_1) + \dots + k_s (x_s - y_s))]$, и мы имеем

$$\begin{aligned} \lim_{M \rightarrow \infty} \int_0^1 |S(N, x)|^2 dx &= \frac{1}{N} + \frac{N(N-1)}{N^2} \int_0^1 dx_1 \dots \int_0^1 dx_s \int_0^1 dy_1 \dots \\ &\dots \int_0^1 dy_s \exp [2\pi i (k_1 (x_1 - y_1) + \dots + k_s (x_s - y_s))]. \end{aligned}$$

Откуда в силу равенства нулю многократного интеграла следует равенство

$$\lim_{M \rightarrow \infty} \int_0^1 |S(N, x)|^2 dx = \frac{1}{N}.$$

Далее, буквально воспроизводя рассуждения, которые имели место при доказательстве теоремы 1.4, получаем

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \int_0^1 |S(N, x)|^2 = 0 \quad \text{для почти всех } x \text{ из } (0, 1).$$

Это и означает асимптотическое по M выполнение критерия Вейля для s -мерных векторов $Y_n = (\alpha_{l_1^n}, \dots, \alpha_{l_s^n})$. $\square$

Замечание. Легко понять, что из асимптотического по M выполнения критерия Вейля не следует асимптотическая независимость вполне равномерная распределенность последовательности. Более точно: имеет место равенство (1.2.11), если в нем X_k заменить на Y_k , но оно может не выполняться для некоторых интегрируемых по Риману функций f . Это справедливо, если в равенстве (1.2.13) сходимость к пределу при $M \rightarrow \infty$ будет достаточно медленной.

В рамках нашей теории мы можем лишь гарантировать, что равенство

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(Y_k) = \int_{D_s} f(X) dX$$

будет выполняться для всех функций f из некоторого класса функций, если этот класс уже, чем класс интегрируемых по Риману функций. Это может быть класс функций с достаточно быстро убывающими коэффициентами Фурье (в частности — все конечные тригонометрические суммы). $\square$

Теорема 1.8. Для каждой интегрируемой в D_s функции $f(X)$ имеет место равенство

$$\lim_{M \rightarrow \infty} \left(\lim_{N \rightarrow \infty} \text{mes} \left\{ \beta : \frac{\sqrt{N}}{\sigma_f} R_N[f, \beta] < y \right\} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y \exp\left(-\frac{u^2}{2}\right) du,$$

где

$$R_N[f, \beta] = \int_{D_s} f(X) dX - \frac{1}{N} \sum_{k=1}^N f(X_k) \quad (1.2.14)$$

и

$$\sigma_f^2 = \int_{D_s} f^2(X) dX - \left(\int_{D_s} f(X) dX \right)^2. \quad \square$$

Доказательство следует из того, что при $N \rightarrow \infty$ и больших M величины $\left(\frac{1}{N} \sum_{k=1}^N f(x_k)\right)^l$ близки к $E\left(\frac{1}{N} \sum_{k=1}^N f(\Xi_k)\right)^l$, где Ξ_k — s -мерные векторы с независимыми в совокупности и равномерно распределенными на $[0, 1]$ компонентами. Подробное доказательство с оценками погрешности содержится в [35]. Предельная теорема такого сорта может быть также получена из общих соображений эргодической теории.

Таким образом, дробные доли показательной функции удовлетворяют асимптотически по M многим требованиям, предъявляемым к независимым реализациям случайной величины, равномерно распределенной на $[0, 1]$.

1.3. Псевдослучайные числа

Как уже говорилось, будем называть псевдослучайными числа, получаемые с помощью детерминированного алгоритма и приближенно обладающие свойствами независимых реализаций равномерно распределенной случайной величины. Дополнительным и немаловажным условием является удобная реализуемость алгоритма с помощью компьютера. Существуют несколько так называемых датчиков случайных (точнее, псевдослучайных) чисел — алгоритмов, которые многократно проверялись с помощью статистических тестов и изучались теоретически. Мы не ставим своей целью проводить анализ состояния проблемы и давать обзор обширной литературы в этой области. Для нас важно показать принципиальную возможность приближенной имитации случайности.

Вычисление дробных долей показательной функции могло бы решить нашу задачу, но соответствующий алгоритм требует, вообще говоря, выполнения операций над вещественными числами — бесконечнозначными дробями — и нереализуем на компьютере. Близким по форме к нему является алгоритм, предложенный Лемером [21]. Как мы увидим далее, между упомянутыми алгоритмами имеется более глубокая связь, чем это принято считать [24]. Остановимся подробно на частном случае алгоритма (датчика), основанного на методе сравнений (иногда в русском переводе “конгруэнтного датчика”). Этот датчик называют мультипликативным. Алгоритм имеет вид

$$\begin{aligned} Z_{n+1} &\equiv MZ_n \pmod{P}, \\ \alpha_{n+1} &= Z_{n+1}/P \end{aligned} \tag{1.3.1}$$

где заданы целые числа $M > 1$, $P > 1$, Z_0 . Не умаляя общности, можно считать, что $M < P$ и $Z_0 < P$. В общем случае линейный метод сравнений для получения псевдослучайных чисел использует сравнение

$$Z_{n+1} \equiv a_0 Z_n + a_1 Z_{n-1} + \dots + a_k Z_{n-k} + b \pmod{P}.$$

Заметим, что первое из равенств (1.3.1) (сравнение по модулю P) есть операция, которая может быть просто выполнена на компьютере (точно, без округлений) как операция над целыми числами. Что касается второго, то при делении происходит округление, кроме некоторых специальных случаев.

Специальным случаем при вычислениях в двоичной системе счисления является $P = 2^m$, и алгоритм состоит в том, что мы перемножаем два целых числа Z_n и M , представленных в разрядной сетке с m разрядами. Результат будет иметь $2m$ разрядов. Последние m разрядов считают ненормированной мантиссой числа α_{n+1} , порядок которого полагают равным нулю.

Хотя случай $P \neq 2^m$ (особенно для простого P) имеет ряд преимуществ (см. [53]), мы будем рассматривать $P = 2^m$, так как это удобно в теоретическом отношении.

Изучим основные свойства датчика (1.3.1).

Лемма 1.1. Для последовательности (1.3.1) целых чисел Z_n для каждого n существует такое натуральное $L' \geq 1$, что $Z_{n+L'} = Z_n$. $\square$

Утверждение леммы следует из того, что число различных Z_n конечно (не превосходит P). Число L' не является единственным числом, обладающим свойством $Z_{n+L'} = Z_n$, $L' \geq 1$. Наименьшее из L' , обладающее этим свойством, называют периодом последовательности (1.3.1).

Очевидно, что для каждого L' справедливо равенство $L' = cL$, где c — натуральное число.

Особый интерес представляет последовательность, имеющая максимальный период.

Исследовать состав последовательности (1.3.1) и свойства ее периодичности позволяют следующие далее леммы. Мы предполагаем, что общий наибольший делитель x_0 и $M - (x_0, M)$ равен 1. Будем считать x_0 фиксированным, и при фиксированном M обозначим последовательность (1.3.1) как $Z(P)$, а длину ее периода как $L(P)$.

Лемма 1.2. Пусть P_1 и P_2 целые положительные, $(P_1, P_2) = 1$. Тогда $L(P_1 P_2) = \text{ОНК}(L(P_1), L(P_2))$, где $\text{ОНК}(m, n)$ означает общее наименьшее кратное натуральных чисел m и n .

Доказательство. Как и ранее $Z_{n+1} = MZ_n \pmod{P}$, $Y_0 = Z_0$ и $Y_{n+1} = MY_n$. Для $i < \min(L(P_1), L(P_2))$ имеем

$$\begin{aligned} Y_{i+jL(P_1)} &= Z_i + C_{i,j}P_1, \\ Y_{i+lL(P_2)} &= Z_i + C_{i,l}P_2, \end{aligned} \quad (1.3.2)$$

где $C_{i,j}, C_{i,l}, i, j, l$ — целые положительные числа. Но также

$$Y_{i+L(P_1 \cdot P_2)} = Z_i + C_i P_1 \cdot P_2, \quad C_i — \text{целое число.} \quad (1.3.3)$$

Из определения периода последовательности $Z(P_1 P_2)$ следует, что $L(P_1 P_2)$ есть наименьшее целое, при котором выполняется (1.3.2), (1.3.3), то есть $L(P_1 P_2) = \text{ОНК}(L(P_1), L(P_2))$. $\square$

Из доказанной леммы легко получить

Следствие. Если $P = \prod_{i=1}^s p_i^{n_i}$ есть разложение числа P на простые множители, то

$$L(P) = \text{ОНК}(L(p_1^{n_1}), \dots, L(p_s^{n_s})). \quad \square$$

Следующие две леммы позволяют исследовать важный для нас случай наличия кратных сомножителей.

Лемма 1.3. Если последовательность $Z(p^m)$ имеет период, длина которого равна $L(p^m)$ (p — простое, m — натуральное числа), то длина периода последовательности $Z(p^{m+1})$ либо равна $L(p^m)$, либо в p раз его больше.

Доказательство. Пусть Y_{n+1} определяется, как и выше, соотношением

$$Y_{n+1} = MY_n, \quad Y_0 = Z_0, \quad (1.3.4)$$

тогда

$$Y_{j+L(p^m)} = Z_j + C_{1,j}p^m.$$

Причем $C_{1,j}$ можно считать неотрицательным числом. Если $C_{1,1}$ делится на p , то $L(p^{m+1}) = L(p^m)$. В противном случае найдется натуральное r , для которого $C_{1,1} \equiv r \pmod{p}$ и, следовательно, $C_{1,j} \equiv jr \pmod{p}$. Так как p — простое, то $jr \not\equiv 0 \pmod{p}$ при $j = 1, 2, \dots, p-1$, и $jr \equiv 0 \pmod{p}$ при $j = p$. Отсюда следует $L(p^{m+1}) = pL(p^m)$, что и завершает доказательство. $\square$

Лемма 1.4. Если функция $X(P)$ такова, что $L(p^{m+1}) = pL(p^m)$, то $L(p^{l+m}) = p^l L(p^m)$, где l и m — натуральные числа.

Доказательство. Сохраняя прежние обозначения, при $l = 2$ имеем $C_{i,p} \equiv pr \pmod{p^r}$. Следовательно, $C_{i,jp} \equiv jpr \pmod{p^r}$. Отсюда, как и ранее, следует $jpr \equiv 0 \pmod{p^r}$, $j = 1, 2, \dots, p-1$, в то время как $ppr \equiv 0 \pmod{p^r}$. Доказательство для $l > 2$ может быть легко проведено по индукции. $\square$

Как уже было сказано, мы рассматриваем подробно лишь случай $P = 2^m$, хотя результат для P , равного степени любого простого числа, получается автоматически. Подробные исследования общего случая см. в [17], [53]. Очевидно, что для нашей последовательности

$$Z_{n+1} \equiv MZ_n \pmod{2^m}, \quad n = 0, 1, \dots \quad (1.3.5)$$

справедлива следующая теорема.

Теорема 1.9. Максимальная длина периода последовательности

$$x_{n+1} \equiv Mx_n \pmod{2^m}$$

при $m \geq 3$ есть 2^{m-2} и достигается при $M \equiv 3 \pmod{8}$ или $M \equiv 5 \pmod{8}$ для всех нечетных x_0 .

Доказательство. Следует из леммы 1.4 и таблицы, содержащей результаты непосредственных вычислений при $m = 1, 2, 3, 4$ и различных M , если заметить, что случай $M > 2^s$ при $P = 2^s$ сводится к случаю $M_1 < 2^s$, где $M \equiv M_1 \pmod{2^s}$ (s — натуральное число):

m	M	Состав последовательности	Длина периода
1	1	1, 1, 1, ...	1
	1	1, 1, 1, ...	1
2	3	1, 3, 1, ...	2
	1	1, 1, 1, ...	1
3	3	1, 3 или 5, 7	2
	5	1, 5 или 3, 7	2
	7	1, 7 или 3, 5	2
4	1	1, 1, 1, ...	1
	3	1, 3, 9, 11 или 5, 15, 13, 7	4
	5	1, 5, 9, 13 или 3, 15, 11, 7	4
	7	1, 7 или 3, 5 или 9, 15 или 11, 13	2
	9	1, 9 или 3, 11 или 5, 13, или 7, 15	2
	11	1, 11, 9, 3 или 5, 7, 13, 15	4
	13	1, 13, 9, 5 или 3, 7, 11, 15	4
	15	1, 15 или 3, 13 или 5, 11 или 7, 9	4

Кроме того, легко видеть, что увеличение периода в два раза с увеличением m от 4 до 5 достигается за счет появления в последовательности новых членов, каждый из которых сравним с одним и только одним из старых по модулю 8. В общем случае при переходе от m к $m+1$ добавляются новые члены, каждый из которых сравним с одним и только одним из старых по модулю 2^m . Это полностью доказывает теорему и приводит к следствию, определяющему состав последовательности с максимальным периодом в рассматриваемом случае. $\square$

Следствие. Справедливы следующие утверждения:

- Если $M \equiv 3 \pmod{8}$ и $x_0 \equiv 1, 3, 9, 11 \pmod{16}$, то последовательность вида $x_{n+1} \equiv Mx_n \pmod{2^m}$ состоит из чисел вида $8k+1$ и $8k+3$, $k = 0, 1, \dots, 2^{m-3} - 1$.
- Если $M \equiv 3 \pmod{8}$ и $x_0 \equiv 5, 7, 13, 15 \pmod{16}$, то эта последовательность состоит из чисел вида $8k+5$ и $8k+7$, $k = 0, 1, \dots, 2^{m-3} - 1$.
- Если $M \equiv 5 \pmod{8}$ и $x_0 \equiv 1 \pmod{4}$, то эта последовательность состоит из чисел вида $4k+1$, $k = 0, 1, \dots, 2^{m-2} - 1$.
- Если $M \equiv 5 \pmod{8}$ и $x_0 \equiv 3 \pmod{4}$, то эта последовательность состоит из чисел вида $4k+3$, $k = 0, 1, \dots, 2^{m-2} - 1$. $\square$

Заметим без доказательства, что для смешанного метода, то есть для последовательности вида

$$x_{n+1} \equiv Mx_n + b \pmod{2^m}, \quad (1.3.6)$$

максимальный период есть 2^m , и последовательность естественным образом состоит из всех чисел от 0 до $2^m - 1$.

Обсудим связь между мультипликативным датчиком (1.3.1) и изучавшимися нами ранее последовательностями дробных долей показательной функции. Можно заметить следующее.

- (1) Соотношение (1.3.1), определяющее датчик псевдослучайных чисел, может быть записано в форме

$$\alpha_{n+1} = \{M\alpha_n\}, \quad \alpha_0 = z_0/2^m, \quad (1.3.7)$$

где z_0 — целое положительное, $z_0 < 2^m$. При явном сходстве (1.3.7) и (1.2.9) очевидно, что в (1.2.9) ε_0 должно быть иррациональным. При рациональном α_0 последовательность (1.3.7) имеет период, и для любой $f(x)$, интегрируемой по Риману, среднее $\frac{1}{N} \sum_{k=1}^N f(x) dx$ не может иметь пределом интеграл $\int_0^1 f(x) dx$ при $m \rightarrow \infty$.

- (2) Пусть M_1 и M_2 — два различных мультипликатора, порождающих одинаковые множества чисел (например, $M_1 \equiv 3 \pmod{8}$)

и $M_2 \equiv 5 \pmod{8}$, а $z_0 \equiv 3 \pmod{4}$) (см. теорему 1.9). Если x принадлежит этому множеству чисел, а разность $|M_1 - M_2| \pmod{2^m}$ велика, то, очевидно, $\{M_1x\}$ и $\{M_2x\}$ могут различаться очень сильно. Это значит, что последовательности, порождаемые M_1 и M_2 , могут не иметь ничего общего, кроме, может быть, распределения. Вопрос о распределении последовательностей, порождаемых датчиками псевдослучайных чисел, мы обсудим далее, принимая во внимание

Определение 1.6. Функцией распределения конечного множества вещественных чисел $a_1, a_2, \dots, a_N$ называется функция (аналог эмпирической функции распределения)

$$\hat{F}_N(x) = \frac{1}{N} \sum_{k=1}^N \chi_{(-\infty, x]}(a_k) = \frac{1}{N} \times N_k,$$

N_k — количество чисел a_k , не превосходящих x , $k = 1, \dots, N$. Функцией распределения последовательности $a_1, a_2, \dots, a_N$ называется $\lim_{N \rightarrow \infty} \hat{F}_N(x)$, если этот предел существует.

- (3) Пусть датчик псевдослучайных чисел порождает последовательность $\alpha_1, \alpha_2, \dots$ рациональных чисел, имеющую период 2^μ , где $\mu = m-2$ для мультипликативного датчика и $\mu = m$ — для смешанного, определяемого (1.3.6). Как и ранее, предполагаем, что функция $f(X)$, $X = (x_1, \dots, x_s)$, интегрируема по Риману в единичном s -мерном гиперкубе. Образует векторы $X_k^{(s)} = (\alpha_{(k-1)s+1}, \dots, \alpha_{ks})$, $k = 1, 2, \dots$, и рассмотрим предел

$$\sum[s, f] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(X_k^{(s)}). \quad (1.3.8)$$

Так как исходная последовательность имеет период 2^μ , то пределом будет конечная сумма, число слагаемых которой зависит от s . При $s = 2^q$, $q < \mu$, число слагаемых будет $2^{\mu-q}$. В случае $s = r \cdot 2^q$, где r — нечетное число, а $q \geq 0$, имеем

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(X_k^{(s)}) &= \\ &= \frac{1}{2^{\mu-q} \cdot r} \sum_{k=1}^{2^{\mu-q}} \sum_{l=1}^r f(\alpha_{(k-1)s+1+2^q(l-1)}, \dots, \alpha_{ks+2^q(l-1)}). \end{aligned}$$

Обозначая $\alpha_{(k-1)s+1+2^q(l-1)}$ через $x_{k,l}$, получим для мультипликативного датчика

$$\sum [s, f] = \frac{1}{2^{\mu-q} \cdot r} \sum_{k,l} f(x_{k,l}, \{Mx_{k,l}\}, \dots, \{M^{s-1}x_{k,l}\}), \quad (1.3.9)$$

где множество $x_{k,l}$ состоит из чисел вида, указанного в следствии к теореме 1.9, разреженных в 2^q раз и деленных на $2^{\mu-q} \cdot r$. Так, если $M \equiv 5 \pmod{8}$ и $x_0 \equiv 3 \pmod{4}$, то упомянутое множество состоит из чисел вида

$$\frac{4k+3}{2^{m-2}} \cdot 2^q = \frac{4k+3}{2^{m-q-2}}, \quad k = 0, 1, \dots, 2^{m-2} - 1.$$

В любом случае мы имеем правильную решетку и, очевидно,

$$\lim_{m \rightarrow \infty} \sum [s, f] = \int_0^1 f(x, \{Mx\}, \dots, \{M^{s-1}x\}) dx$$

для всякой интегрируемой по Риману функции $f(X)$.

При $s=1$ равенство (1.3.9) означает, что распределение точек последовательности при $m \rightarrow \infty$ слабо сходится к равномерному. При $s > 1$ и фиксированном M можно лишь утверждать, что распределение векторов $(\alpha_{(k-1)s+1}, \dots, \alpha_{ks})$ сходится в слабом смысле к распределению векторов, составленных аналогичным образом из последовательности дробных долей показательной функции $x, \{Mx\}, \{M^2x\}, \dots$ для почти всех x из $(0, 1)$.

Равенство (1.3.9) также можно записать в виде

$$\lim_{m \rightarrow \infty} \sum [s, f] = \int_0^1 \varphi_M(x) dx \quad (1.3.10)$$

и применить теорему Сю. Это приведет к желаемому пределу в виде s -кратного интеграла. Однако мы во всех случаях рассматривали повторные пределы. Реально же мы имеем дело с последовательностями конечных M и P . Если не удастся доказать существование двойного предела по M и P , необходимо, по крайней мере, доказать существование последовательностей (M, P) , таких что этот предел существует.

Анализ, как и ранее, достаточно провести при произвольном фиксированном s для гармоник вида $f(\mathbf{k}, X) = \exp\{2\pi i(k_1x_1 + \dots + k_sx_s)\}$, где $\mathbf{k} = (k_1, \dots, k_s)$. В случае рассмотренного нами датчика имеем из (1.3.9)

$$\sum [s, f(\mathbf{k}, X)] = \frac{1}{2^{n-q}r} \sum_{k,l} \exp\{2\pi i(k_1 + k_2M + \dots + k_sM^{s-1})x_{k,l}\}.$$

Если выполняется сравнение

$$k_1 + k_2 M + \dots + k_s M^{s-1} \equiv 0 \pmod{P}, \quad (1.3.11)$$

то $\sum[s, f(\mathbf{k}, X)] = 1$, а предел при $M \rightarrow \infty$ для всякой такой суммы должен быть нулем согласно (1.3.10).

При фиксированном P возможен, вообще говоря, такой выбор последовательности чисел $M \rightarrow \infty$, при которой для каждого M существуют “плохие” $\mathbf{k}$, то есть выполняется сравнение (1.3.11). И возникает вопрос: всегда ли можно выбрать “хорошую” последовательность?

Нужная последовательность пар (M, P) может быть построена.

Действительно, при произвольном фиксированном s обозначим K — множество векторов $\mathbf{k} = (k_1, \dots, k_s)$, для которых выполняется (1.3.11) и $\|\mathbf{k}\| \leq L$.

Справедлива следующая лемма.

Лемма 1.5. Для заданного L существуют такие M и P , что множество K пусто.

Доказательство. Сравнение (1.3.11) равносильно множеству равенств

$$k_1 + k_2 M + \dots + k_s M^{s-1} = lP, \quad l = 0, \pm 1, \pm 2, \dots \quad (1.3.12)$$

В силу эквивалентности норм конечномерных векторов существует константа $L' > 0$, такая что $\max_i |k_i| \leq L'$, и при достаточно больших M значение многочлена относительно M в левой части (1.3.12) определяется членом $k_s M^{s-1}$. Таким образом, при $l = 0$ найдется такое $M_0 > 0$, что при $M > M_0$ многочлен не будет менять знак. С другой стороны, этот член ограничен величиной $L' M^{s-1}$, и при заданном s найдется такое достаточно большое $P = P(M)$, что многочлен

$$k_s M^{s-1} + \dots + k_2 M + k_1 - lP$$

не имеет вещественных корней при любом значении $l = \pm 1, \dots$ (напомним, что все $|k_i|$ ограничены!). Лемма доказана. $\square$

Таким образом, конкретный мультипликативный датчик не решает задачу имитации случайностей даже в той ограниченной постановке, которую мы сформулировали в начале главы. Однако последовательность датчиков с возрастающими M и P могут ее решить для любого заданного L .

Изложенные выше соображения были призваны выяснить причины, по которым оказалось возможным широко использовать мультипликативный датчик при решении прикладных задач. Вопросы выбора мультипликатора подробно изложены в доступной литературе (см.

библиографический комментарий к главе). Стараются выбрать M таким, чтобы L , фигурирующее в лемме 1.5, было наибольшим. Обычно используется евклидова норма вектора $\mathbf{k}$.

Принято (см. [3], [53]) числа

$$V_s(M) = \max_{k_1, \dots, k_s} ((k_1^2 + k_2^2 + \dots + k_s^2)^{1/2} : \\ k_1 + k_2 M + \dots + k_s M^{s-1} \not\equiv 0 \pmod{P})$$

считать характеристиками качества датчика псевдослучайных чисел. Целесообразность введения такого критерия мы объяснили ранее. Величину $V_s(M)$ называют волновым числом.

1.4. Библиографическое дополнение к главе 1

Теория равномерно распределенных в смысле Вейля последовательностей превосходно изложена в книге Кейперса и Нидеррейтера [52]. Проблема арифметического моделирования случайности восходит к работам Колмогорова и Линника [10], [7].

С методом Монте-Карло непосредственно связана работа Франклина [10], посвященная моделированию случайности.

В главе 1 предпринята попытка по возможности элементарного подхода к проблеме. Теорема 1.4 заимствована из [52]. Понятие s -независимо равномерно распределенных последовательностей менее изящно, чем понятие s -равномерно распределенных последовательностей, но отражает специфику метода Монте-Карло. Теорема 1.5 есть перефразировка аналогичного утверждения из [10].

Предельная (асимптотическая по M) теорема доказана впервые автором (см. книгу [35]).

Таким образом, удастся продемонстрировать, что дробные доли показательной функции приближенно обладают многими свойствами “настоящих” случайных чисел.

Переход к свойствам реального мультипликативного датчика псевдослучайных чисел вызывает дополнительные трудности. Если одномерное распределение последовательности Z_n/P , $Z_{n+1} \equiv MZ_n \pmod{P}$ сходится к равномерному при $P \rightarrow \infty$ и любом целом $M > 1$, то распределение последовательности s -независимых векторов требует для сходимости к равномерному распределению в единичном s -мерном гиперкубе при $P \rightarrow \infty$, как мы видели, специального выбора последовательности чисел $M \rightarrow \infty$. Такой выбор может быть связан со свойствами так называемых волновых чисел. Существуют специальные методы их численной оценки. По этому поводу см., например, [3], [53]. Разными авторами было потрачено много

машинного времени для отбора мультипликаторов с хорошими свойствами. Примеры таких мультипликаторов можно найти в [8], [11].

Некруткиным рассматривалась модель мультипликативного датчика, в которой выбор начального значения осуществляется случайно. Это позволяет привлечь к исследованию средства теории вероятностей (аппарат характеристических функций и др.). При стремлении к бесконечности P , M и волновых чисел приходим в пределе к независимым реализациям равномерно распределенных случайных чисел (см. [11], [60]). В этих же работах исследуются некоторые мультипликаторы специального вида. Приводятся примеры “плохих” последовательностей мультипликаторов, для которых при $M \rightarrow \infty$ предельное распределение не будет равномерным.

Рекуррентные соотношения вида $x_n \equiv g(x_{n-1}) \pmod{P}$, где g — нелинейная функция, рассматривались Нидеррейтером [24]. Нелинейность была призвана нарушить так называемую решетчатую структуру последовательностей, возникающую в случае нелинейной функции g . В линейном случае (см. [23], а также [24]) векторы, образованные из элементов последовательности, принадлежат некоторому набору гиперплоскостей. Геометрическая интерпретация этого факта может быть полезной.

Отметим, что отображение точки $x \in (0, 1)$ в точку $(x, Mx, \dots, m^{s-1}x)$ s -мерного куба при $M \rightarrow \infty$ приближенно осуществляет отображение отрезка на гиперкуб — гиперплоскости все более плотно наполняют его. Такого рода отображения, если они просты в алгоритмическом отношении, могут быть связаны с псевдослучайными числами. Покрытия гиперкуба гиперповерхностями типа кривой Пеано используются в задачах глобального поиска экстремума [73].

В последние годы значительное развитие получили методы, использующие рекуррентные связи между разрядами последовательно получаемых псевдослучайных чисел. Эта тема изложена в работе Таусворта ([26], русский перевод [74]).

Основы общей теории рекуррентных поразрядных соотношений обсуждаются в неоднократно упоминаемой книге Нидеррейтера [24]. Важным развитием исследований в этом направлении явилась работа [21], на основе которой был создан датчик “Мерсенна Твистер”, имеющий период $2^{19937} - 1$ и обеспечивающий s -равномерность распределений вплоть до $s = 623$. Датчик с такими параметрами удовлетворяет практически всем известным потребностям. В частности, на его базе легко создавать независимые подпоследовательности, используемые при параллельных вычислениях.

Глава 2

Моделирование случайных величин и процессов

2.1. Моделирование случайных величин с заданным законом распределения

Далее мы рассмотрим некоторые методы моделирования случайных величин с заданным законом распределения (отличным от равномерного на промежутке $[0, 1]$). Задача моделирования упомянутых случайных величин имеет очень большое прикладное значение. Ей посвящена обширная литература. Мы сошлемся на монографию Девроя [4], а также на важную работу Кнута и Яо [54]. В этих трудах имеется много ссылок на работы по рассматриваемому вопросу.

Остановимся только на общих методах моделирования. Конкретные распределения будут фигурировать лишь в качестве примеров применения общих методов.

Будем различать два подхода:

- (1) исходной для моделирования нужной случайной величины является случайная величина, равномерно распределенная на $[0, 1]$ (предполагается, что она задается с необходимым для наших целей числом двоичных разрядов);
- (2) исходной является последовательность случайных битов, на основе которой вычисляется последовательность битов двоичного представления моделируемой случайной величины.

Условно можно считать, что для первого подхода используется мультипликативный датчик (1.3.1), а для второго — датчик случайных битов Таусворта (1.3.12). Сначала будем рассматривать первую задачу, считая исходными независимые реализации α случайной величины, равномерно распределенной на $[0, 1]$.

2.1.1. Формула обращения

Хорошо известна формула обращения, дающая принципиальное решение задачи вычисления реализации ξ случайной величины с заданной непрерывной функцией распределения F . Реализация ξ выражается через реализацию равномерно распределенной на $[0, 1]$ случайной величины α . Справедлива

Теорема 2.1. Пусть F — непрерывная функция распределения на $\mathbf{R}^1$ и пусть обратная к ней определена формулой

$$F^{-1}(x) = \inf\{t : F(t) = x; 0 < x < 1\}, \quad (2.1.1)$$

тогда если α — равномерно распределенная на $[0, 1]$ случайная величина, то $F^{-1}(\alpha)$ имеет своей функцией распределения F .

Доказательство. Пусть $\xi = F^{-1}(\alpha)$. Поскольку $F_\xi(x) = P(\xi \leq x)$, то имеем

$$\begin{aligned} P(\xi \leq x) &= P(F^{-1}(\alpha) \leq x) = \\ &= P(\inf\{t : F(t) = \alpha\} \leq x) = P(\alpha \leq F(x)). \end{aligned}$$

Далее, так как $0 \leq F(x) \leq 1$, а α равномерно распределена на $[0, 1]$, то $P(\alpha \leq F(x)) = F(x)$, то есть $F_\xi(x) = F(x)$, что и доказывает теорему. $\square$

Замечание. Легко доказать и обратное утверждение: если ξ — случайная величина с непрерывной функцией распределения $F(x)$, то $F(\xi)$ равномерно распределена на промежутке $[0, 1]$. Действительно, для $x \in (0, 1)$ имеем $P(F(\xi) \leq x) = P(\xi \leq F^{-1}(x)) = F(F^{-1}(x)) = x$, что и доказывает последнее утверждение. $\square$

Непосредственное использование формулы обращения в вычислительных процедурах, как правило, малоэффективно. Нужно, чтобы F имела достаточно простое аналитическое выражение. Рассмотрим примеры.

Примеры. (1) Экспоненциальное распределение (λ — параметр):

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0, \quad \lambda > 0,$$

$$F^{-1}(\alpha) = -\frac{1}{\lambda} \ln(1 - \alpha).$$

Или с учетом того, что $1 - \alpha$ распределено равномерно на $[0, 1]$ (то есть как и α), имеем

$$\xi = -\frac{1}{\lambda} \ln \alpha. \quad (2.1.2)$$

(2) Распределение Парето (a и b — параметры):

$$F(x) = 1 - \left(\frac{b}{x}\right)^a, \quad x \geq b > 0,$$

$$F^{-1}(\alpha) = \frac{b}{(1 - \alpha)^{1/a}} \quad \text{и} \quad \xi = \frac{b}{\alpha^{1/a}}. \quad (2.1.3)$$

□

Легко видеть, однако, что весьма популярное нормальное распределение не допускает простого вычисления $F^{-1}(\alpha)$. Решение уравнения

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-u^2/2} du = \alpha \quad (2.1.4)$$

требует либо использования сложных асимптотических рядов, либо запоминания достаточно подробной таблицы функции

$$\operatorname{erf}(x) = \sqrt{\frac{2}{\pi}} \int_0^x e^{-u^2/2} du.$$

С ее помощью путем интерполирования можно достаточно точно находить ξ , соответствующее данному α в (2.1.4).

Здесь мы видим, что существует принципиальная разница между

- (1) задачами, которые допускают предварительную алгоритмическую подготовку: например, составление таблиц обратной функции $F^{-1}(x)$, построение полиномиальных (обычно кусочно-полиномиальных) ее аппроксимаций и т. п.;
- (2) задачами, где значительные затраты на предварительные вычисления не оправданы и требуется “не слишком плохой” алгоритм моделирования, позволяющий решить (одноразовую) задачу с разумными затратами времени компьютера.

Разумеется, задача моделирования нормального распределения есть, как правило, задача (1). Имеется обширная библиография, посвященная специально этому распределению. Здесь в связи с формулой обращения полезно принять во внимание тот любопытный факт, что для получения одной реализации нормально распределенной ξ формула обращения непригодна, но получить сразу две точно независимые реализации с ее помощью сравнительно просто.

Пример. Действительно, пусть ξ и η независимы и распределены по стандартному нормальному закону. Плотность их совместного распределения есть

$$\frac{1}{2\pi} e^{-(x^2+y^2)/2}.$$

Переход к полярным координатам $\xi = \rho \cos \psi$, $\eta = \rho \sin \psi$ дает для плотности совместного распределения случайных величин ρ и ψ выражение

$$\frac{1}{2\pi} r e^{-r^2/2}.$$

Мы видим, что случайный угол ψ равномерно распределен на окружности (с плотностью $\frac{1}{2\pi}$), а ρ имеет плотность $r e^{-r^2/2}$, $r \geq 0$, к которой легко применяется формула обращения. Окончательно получим

$$\xi = \sqrt{-2 \ln \alpha_1} \cos 2\pi \alpha_2, \quad \eta = \sqrt{-2 \ln \alpha_1} \sin 2\pi \alpha_2. \quad (2.1.5)$$

□

Можно сказать, что задачи эффективного моделирования случайной величины — это не просто задачи выбора удобного метода решения уравнения $F(\xi) = \alpha$. Они обладают несомненным своеобразием и требуют развития специальных методов.

Пример. Алгоритм моделирования дискретной случайной величины, принимающей значения $a_1, \dots, a_n, \dots$ соответственно с вероятностями $p_1, \dots, p_n, \dots$, формально весьма схож с алгоритмом решения уравнения $F(x) = \alpha$. Введение несколько более общего определения обратной функции может привести нас прямо к этой формуле, но очень простые рассуждения позволяют решить задачу непосредственно. Итак, пусть дискретное распределение задано таблицей

a_1	a_2	$\dots$	a_n	$\dots$
p_1	p_2	$\dots$	p_n	$\dots$

причем $\sum_{i=1}^{\infty} p_i = 1$, a_i различны и, не умаляя общности, предполагается, что $p_i > 0$, $i = 1, 2, \dots$. Обозначим $S_n = \sum_{i=1}^n p_i$. Заметим, что из определения равномерного распределения следует, что

$$P(S_{n-1} \leq \alpha < S_n) = p_n, \quad n = 2, \dots \quad (2.1.6)$$

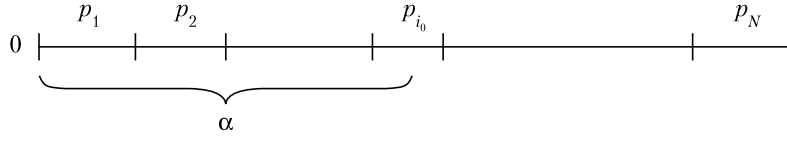


Рис. 2.1

Отсюда вытекает, что случайный номер i_0 величин $a_1, \dots, a_n, \dots$, соответствующий данному α , может быть определен из неравенства

$$S_{i_0-1} \leq \alpha < S_{i_0}, \quad (2.1.7)$$

то есть $0 \leq \alpha - S_{i_0-1} < p_{i_0}$, см. рис. 2.1, а также ср. (2.1.1). $\square$

Замечание. Алгоритм нахождения i_0 есть алгоритм поиска в таблице. В случае когда таблица конечна (имеет размер N), нужное i_0 можно найти по крайней мере за $\log_2 N$ операций. $\square$

Алгоритм. Необходимо предварительно вычислить суммы S_n . Храним массив $S_n, n = 1, \dots, N$, и текущие данные: $\langle \text{начало} \rangle, S(\text{нач.}), \langle \text{конец} \rangle, S(\text{кон.}), r$.

1. $\langle \text{начало} \rangle := 0, \quad \langle \text{конец} \rangle := N,$ — начальные засылки.
 $S(\text{нач.}) = 0, \quad S(\text{кон.}) = 1$
2. α . Вычисляем реализацию равномерно распределенной случайной величины.
3. $r := \lfloor \frac{\langle \text{конец} \rangle - \langle \text{начало} \rangle}{2} \rfloor$. Здесь $\lfloor a \rfloor$ означает целую часть числа a .
4. $S(r) < \alpha$.
 Если да, то 5; если нет, то 6.
5. $\langle \text{начало} \rangle + r; \quad S(\text{нач.}) := S(\text{нач.}) + S(r)$.
6. $\langle \text{конец} \rangle := \langle \text{конец} \rangle - r; \quad S(\text{кон.}) := S(\text{кон.}) - S(r)$.
7. $\langle \text{конец} \rangle - \langle \text{начало} \rangle \leq 1$.
 Если нет, то 3; если да, то r есть результат работы алгоритма.

Если N бесконечно, то используем другие методы, например метод композиции (см. далее).

2.1.2. Использование аппарата условных вероятностей

Далее будем рассматривать задачу моделирования случайного вектора $\Xi = (\xi_1, \dots, \xi_s)$, предполагая существование плотности $f(x_1, \dots, x_s)$

совместного распределения его компонент на $\mathbf{R}^s$. Отметим, что это предположение делается скорее для упрощения обозначений, чем по существу.

Компоненты реализации случайного вектора Ξ могут быть получены путем последовательного моделирования следующей последовательности плотностей:

$$\begin{aligned}\varphi_1(x_1) &= \int_{-\infty}^{+\infty} dx_2 \dots \int_{-\infty}^{+\infty} dx_s f(x_1, \dots, x_s) \quad \text{для } \xi_1, \\ \varphi_2(x_2/\xi_1) &= \frac{\int_{-\infty}^{+\infty} dx_3 \dots \int_{-\infty}^{+\infty} dx_s f(\xi_1, x_2, \dots, x_s)}{\int_{-\infty}^{+\infty} dx_2 \dots \int_{-\infty}^{+\infty} dx_s f(\xi_1, x_2, \dots, x_s)} \quad \text{для } \xi_2, \quad (2.1.8) \\ \varphi_s(x_s/\xi_1, \dots, \xi_{s-1}) &= \frac{f(\xi_1, \dots, \xi_{s-1}, x_s)}{\int_{-\infty}^{+\infty} f(\xi_1, \dots, \xi_{s-1}, x_s) dx_s} \quad \text{для } \xi_s,\end{aligned}$$

то есть ξ_1 находится путем моделирования $\varphi_1(x)$ — частной (одномерной) плотности ее распределения; и далее, если найдены выборочные значения $\xi_1, \dots, \xi_k$ ($k < s$), строится условная плотность распределения ξ_{k+1} и моделируется соответствующее выборочное значение ξ_{k+1} .

Процедура требует s независимых реализаций равномерно распределенных случайных величин $\alpha_1, \dots, \alpha_s$ и принципиально решает задачу моделирования случайного вектора Ξ . Необходимость вычисления интегралов делает ее еще более сложной, чем в одномерном случае, однако возможность разного рода замен переменных позволяет в отдельных случаях получать от нее практическую пользу.

Разумеется также, что нумерацию переменных x_i можно изменить произвольным образом.

Если непосредственное практическое применение формул (2.1.8) приводит к сложным алгоритмам, то теоретическое использование их или их модификаций может быть весьма плодотворным.

Как мы уже видели, в случае применения формулы обращения для моделирования нормального распределения увеличение размерности может давать неожиданно положительный эффект. Такой прием служит источником эффективных алгоритмов и во многих других случаях.

Пусть ставится задача моделирования случайного вектора $\Xi = (\xi_1, \dots, \xi_s)$ с плотностью $\varphi(x_1, \dots, x_s)$ совместного распределения компонент. Добавим еще одну компоненту ξ_{s+1} так, что плотность

компонент вектора (Ξ, ξ_{s+1}) будет задана формулой

$$f(x_1, \dots, x_s, x_{s+1}) = \begin{cases} 1/C, & 0 \leq x_{s+1} \leq \varphi(x_1, \dots, x_s)C, \quad C > 0, \\ 0, & x_{s+1} > \varphi(x_1, \dots, x_s) \text{ или } x_{s+1} < 0, \end{cases}$$

то есть вектор (Ξ, ξ_{s+1}) равномерно распределен в области, ограниченной поверхностью $x_{s+1} = C\varphi(x)$ и гиперплоскостью $x_{s+1} = 0$.

Пусть $\varphi(x_1, \dots, x_s) \geq 0$ в $\mathbf{R}^s$ и интеграл $\int_{\mathbf{R}^s} \varphi_1(x) dx = C$ конечен.

Из выражения для плотности f и (2.1.8) легко получить, что

– частная (маргинальная) плотность распределения вектора Ξ есть

$$\int dx_{s+1} f(x, x_{s+1}) = \int_0^{C\varphi(x)} 1/C dx_{s+1} = \varphi_2(x); \quad (2.1.9)$$

– условная плотность случайной величины ξ_{s+1} при фиксированном векторе $\Xi = \tilde{\Xi}$ есть

$$\frac{f(\tilde{\Xi}, x_{s+1})}{\int_{\mathbf{R}^1} f(\tilde{\Xi}, x) dx} = \frac{1/C}{1/C \int_0^{C\varphi(\tilde{\Xi})} dx_{s+1}} = \frac{1}{C\varphi(\tilde{\Xi})}. \quad (2.1.10)$$

Из формул (2.1.9), (2.1.10) вытекает следующее правило моделирования случайного вектора, равномерно распределенного в области $A = \{(x_1, x_2, \dots, x_{s+1}) : 0 \leq x_{s+1} \leq C\varphi(x_1, \dots, x_s)\}$:

- Моделируя случайный вектор Ξ с плотностью распределения $\varphi(x)$, получаем значение $\Xi = \tilde{\Xi}$.
- На промежутке $[0, 1/C\varphi(\tilde{\Xi})]$ моделируем случайную величину, равномерно распределенную на этом промежутке. Получаем ее реализацию $\xi_{s+1} = \tilde{\xi}_{s+1}$.

Случайный вектор $(\tilde{\Xi}, \tilde{\xi}_{s+1})$, построенный таким образом, будет равномерно распределен на множестве A в силу (2.1.10).

Из (2.1.9) непосредственно следует

Утверждение. Если случайный вектор (Ξ, ξ_{s+1}) равномерно распределен в области A , то:

- (1) Ξ имеет плотностью распределения $\varphi(x)$;
- (2) при фиксированном Ξ и $\varphi(\Xi) > 0$ случайная величина ξ_{s+1} равномерно распределена на промежутке $[0, C\varphi(\Xi)]$.

Последнее утверждение следует из (2.1.10). $\square$

Алгоритм 2.1. Моделирование случайного вектора (Ξ, ξ_{s+1}) , равномерно распределенного в области

$$D_g = \{(x_1, \dots, x_{s+1}) : 0 \leq x_{s+1} \leq g(x_1, \dots, x_s), |D_g| < \infty\}, g = c\varphi(x).$$

1. Вычисляем реализацию $\hat{\Xi}$ вектора Ξ , распределенного с плотностью $1/|D_g|g(x_1, \dots, x_s)$.
 2. Находим реализацию $\hat{\xi}_{s+1}$ случайной величины, равномерно распределенной на отрезке $[0, g(\hat{\Xi})]$.
- Вектор $(\hat{\Xi}, \hat{\xi}_{s+1})$ есть решение задачи.

Замечание. Если $D_0 \subset D_g$, то реализации вектора (Ξ, ξ_{s+1}) , равномерно распределенного в D_g , очевидно, равномерно распределены в D_0 (условная плотность (Ξ, ξ_{s+1}) при условии $(\Xi, \xi_{s+1}) \in D_0$ есть константа). Это замечание служит обоснованием простейшего варианта метода отбора. $\square$

Алгоритм 2.2. (Метод отбора.) Моделирование равномерно распределенного в D_0 вектора при условии, что мы умеем моделировать равномерно распределенный вектор в $D_1, D_0 \subset D_1$ (например, с помощью алгоритма 2.1, если $D_1 = D_g$ при некоторой g).

1. Моделируем вектор, равномерно распределенный в D_1 , вычисляем его реализацию $(\hat{\Xi}, \hat{\xi}_{s+1})$.
 2. Если $(\hat{\Xi}, \hat{\xi}_{s+1}) \in D_0$, то это есть нужное значение вектора; если иначе, то возвращаемся к 1.
- При условии $|D_0| > 0$ процесс заканчивается с вероятностью 1. Величина $|D_0|/|D_1|$ называется эффективностью метода отбора.

Теперь мы можем сформулировать метод отбора для моделирования случайного вектора с заданной плотностью в его наиболее известной классической форме. Пусть плотность $\varphi(x_1, \dots, x_s)$ ограничена и отлична от нуля на ограниченном множестве G ; $g(x_1, \dots, x_s) \leq M$, а G содержится в параллелепипеде $G \subset \Pi = \{x_1, \dots, x_s : a_i \leq x_i \leq b_i, i = 1, \dots, s\}$.

Алгоритм 2.3. (Классический метод отбора.)

1. Получаем реализацию $\hat{\Xi}$ случайного вектора Ξ , равномерно распределенного в G (например, с помощью алгоритмов 2.1–2.2; моделировать равномерное распределение в Π мы уже умеем).

2. Получаем реализации $\hat{\xi}_{s+1}$ случайной величины ξ_{s+1} , равномерно распределенной на промежутке $[0, M]$.
3. Проверяем неравенство $\hat{\xi}_{s+1} \leq \varphi(\hat{\Xi})$. Если оно выполнено, то $(\hat{\Xi}, \hat{\xi}_{s+1})$ есть нужный нам вектор, а если нет, то возвращаемся к 1.

Эффективность описанного алгоритма есть $1/M|\Pi|$, если для моделирования Ξ также используется метод отбора.

Более общим вариантом метода отбора, не требующим ограниченности φ и области G , является метод мажорант.

Пусть $f(X) \geq \varphi(X)$, $\varphi(X)$ — плотность, подлежащая моделированию. Моделируем равномерное распределение в области

$$D_f = \{X : 0 \leq x_{s+1} \leq f(X)\}.$$

Объем $|D_f|$ предполагается конечным. Очевидно, $|D_f| \geq 1 = |D_\varphi|$, и можно использовать алгоритм 2.2. Это дает

Алгоритм 2.4. (Метод мажорант.)

1. Получаем реализацию $\hat{\Xi}$ случайного вектора, распределенного с плотностью $f(X)/|D_f|$.
2. Получаем реализацию $\hat{\xi}_{s+1}$ случайной величины, равномерно распределенной на промежутке $[0, f(\hat{\Xi})]$.
3. Проверяем условие $\hat{\xi}_{s+1} \leq \varphi(\hat{\Xi})$. Если оно выполнено, то $\hat{\Xi}$ — искомое значение случайного вектора, распределенного с плотностью $\varphi(X)$; если нет, то переходим к п. 1.

Эффективность метода есть $1/|D_f|$.

Примеры. (1) $\varphi(x) = Ce^{-x} \cos^2 x$, $x \in [0, \frac{\pi}{2}]$.

$\int e^{-x} \cos^2 x dx = \frac{e^{-x}}{5} (\sin 2x - \cos^2 x - 2) + \text{const}$. Константа нормировки $C = 1/5(3 - 2e^{-\pi/2})$. Формула обращения $e^{-\xi}(2 + \cos^2 \xi - \sin 2\xi) = 3 - (3 - 2e^{-\pi/2})\alpha$. Вычисление значения ξ требует решения трансцендентного уравнения.

С другой стороны, плотность мажорируется функцией $y = C \cos^2 x$ или $y = Ce^{-x}$, можно выбрать более простой вариант. Моделирование показательного распределения проще, но эффективность отбора выше при $y = C \cos^2 x$. Впрочем, в этом случае можно применить простейший вариант метода отбора. Полезно сравнить эффективности всех методов.

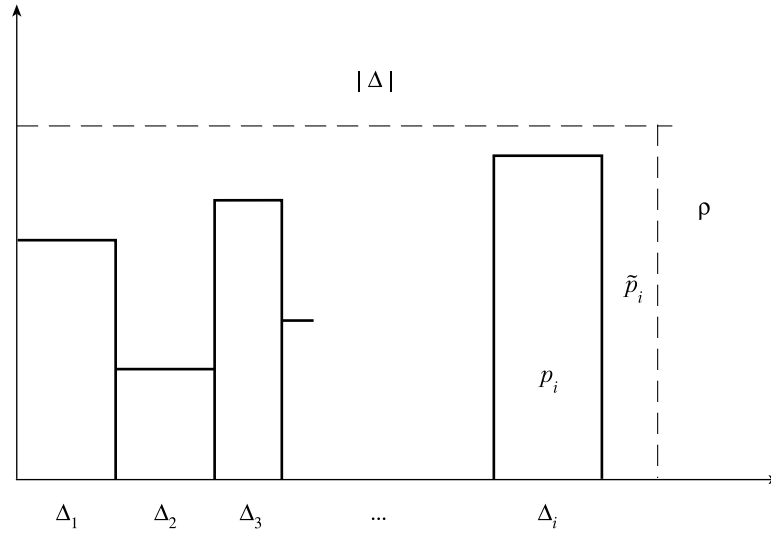


Рис. 2.2

(2) $\varphi(x) = xe^{2-x}/(x-1)^2$, $x \in [2, \infty)$.

Легко проверить, что формула обращения $e^{2-\xi}/(\xi-1) = \alpha$ в этом случае также требует решения трансцендентного уравнения. Здесь необходимо применить метод мажорант. Мажорантой является $y = e^{2-x}$. Промежуток неограничен. $\square$

Вернемся к случаю дискретного распределения. Пусть случайная величина принимает значение n с вероятностью p_n , $\sum_{n=1}^{\infty} p_n = 1$, $p_n \geq 0$. Рассмотренный выше прием увеличения размерности здесь также приводит к интересным результатам. Пусть D — любая область в $\mathbf{R}^2$, разделенная на непересекающиеся подобласти D_n , $\bigcup_{n=1}^{\infty} D_n = D$ такие, что площадь D_n есть p_n . Тогда если мы умеем моделировать равномерно распределенную точку в D , то попадание ее в D_n может интерпретироваться как исход с номером n для заданной дискретной случайной величины.

Это соображение служит обоснованием метода отбора в дискретном случае. Действительно, пусть N — конечное число. Выберем N отрезков Δ_i длиной $|\Delta_i| > 0$ и положим $\tilde{p}_i = p_i/|\Delta_i|$. Отложим отрезки длины $|\Delta_i|$, $i = 1, \dots, N$, на вещественной оси и построим на каждом Δ_i прямоугольник высотой $\tilde{p}_i$ (рис. 2.2), так что площадь i -го прямоугольника есть p_i . Пусть $\rho = \max_i p_i$. В этом случае вся фигура

содержится в прямоугольнике A со сторонами ρ и $|\Delta| = \sum_{i=1}^N \Delta_i$ соответственно.

Теперь моделирование нашей случайной величины производится следующим образом.

Алгоритм 2.5. (Метод отбора в дискретном случае.)

1. Получаем ξ_1 и ξ_2 — две независимые реализации случайных величин, распределенных на $[0, \rho]$ и $[0, |\Delta|]$ соответственно.
2. Если точка (ξ_1, ξ_2) принадлежит прямоугольнику с номером n , то n есть нужная нам реализация. В противном случае переходим к 1.
Легко видеть, что при $|\Delta_i| = p_i$ мы приходим к случаю, рассмотренному ранее, и реализацию ξ_1 можно не получать.

Подобная интерпретация дискретного распределения позволяет построить метод мажорант (N в этом случае не обязано быть конечным). Пусть $H(x)$ — кусочно постоянная функция, равная h_i на Δ_i , $i = 1, 2, \dots, N$, и $h_i \geq p_i$. Предположим, что $H = \sum_{i=1}^N h_i |\Delta_i| < \infty$. Обозначим $A = \{x, y : 0 \leq y \leq H(x)\}$. Предполагаем, что мы умеем моделировать точку, равномерно распределенную в A . Для этого, в частности, достаточно уметь моделировать дискретное распределение $\{h_i/H\}_{i=1}^N$.

Алгоритм 2.6. (Метод мажорант в дискретном случае.)

1. Получаем реализацию r случайной величины с распределением $\{h_i/H\}_{i=1}^N$.
2. Получаем реализацию ξ случайной величины, равномерно распределенной на $[0, h_r]$.
3. Проверяем условие $\xi \leq p_r$. Если да, то r — нужное нам число, если нет, то переходим к 1.

Если все $|\Delta_i|$ равны между собой, то условие п. 3 выглядит следующим образом: $\alpha h_r \leq p_r$, где α — равномерно распределенная случайная величина на $[0, 1]$. Предполагается, что $h_r > 0$, если $p_r > 0$. В этом случае говорят также, что распределение $(p_1, \dots, p_N)$ абсолютно непрерывно по отношению к $(h_1, \dots, h_N)$.

Замечания. (1) (О моделировании случайного вектора с дискретным распределением.) Если каждая из компонент ξ_r случайного вектора $(\xi_1, \dots, \xi_s)$ может принимать целочисленные значения $1, 2, \dots, n_r$, то

совместное распределение его компонент задается таблицей

$$\{P_{i_1, \dots, i_s}; i_r = 1, \dots, n_r; r = 1, \dots, s\}.$$

Не составляет труда перенести описанные ранее приемы на этот случай.

(2) Легко получить обобщения методов отбора, когда в качестве основных распределений выбираются не равномерные, а любые другие заданные распределения. Соответствующие формулы можно найти в [4], [8]. $\square$

Метод композиции

Пусть $F_\eta(y)$ — заданное распределение случайной величины η , а $p(X, y) = P(\Xi = X/\eta = y)$ — условная плотность распределения вектора Ξ при фиксированном η . Тогда плотность распределения Ξ равна

$$\varphi_\Xi(X) = \int_{-\infty}^{+\infty} p(X, y) dF_\eta(y) \quad (2.1.11)$$

и может моделироваться в два этапа:

- (1) Находим реализацию $\hat{\eta}$ случайной величины η с функцией распределения $F_\eta(y)$.
- (2) Моделируем вектор Ξ , имеющий плотность $p(X, \hat{\eta})$.

Несмотря на свою простоту метод композиции очень важен при построении эффективных алгоритмов моделирования.

Примеры. (1) Пусть функция $\psi(x)$ представлена в виде ряда

$$\psi(x) = \sum_{k=0}^{\infty} a_k x^k, \quad a_k \geq 0, \quad x \in [0, 1].$$

Из условия нормировки имеем $\sum_{k=0}^{\infty} a_k/(k+1) = 1$. Представим $\psi(x)$ в форме

$$\psi(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (k+1)x^k,$$

и можно считать k случайным номером, распределение которого есть $p_k = a_k/(k+1)$, k играет роль y , а $p(x, y)$ отождествляется с $(k+1)x^k$.

Алгоритм состоит в выборе случайного номера k_0 в соответствии с распределением $\{a_k/(k+1)\}$, а затем моделировании плотности $(k+1)x^{k_0}$.

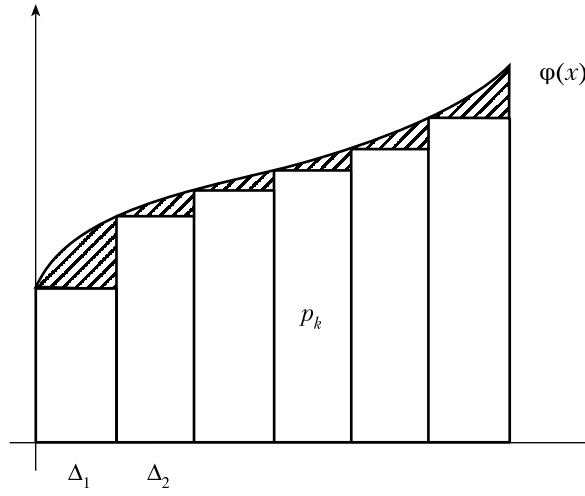


Рис. 2.3

(2) (Более общий.) Плотность $\psi(x)$, $x \in \mathbf{R}^1$, разбивается на сумму плотностей следующим образом. Вещественная ось представляется в виде суммы непересекающихся интервалов Δ_k , $k = 1, \dots, N$. Вычисляются числа

$$p_k = \int_{\Delta_k} \varphi(x) dx$$

и полагается

$$\psi(x) = \sum_{k=1}^N p_k \varphi_k(x), \quad \varphi_k(x) = \begin{cases} 1/p_k \cdot \varphi(x), & x \in \Delta_k, \\ 0, & x \notin \Delta_k \end{cases} \quad (2.1.12)$$

(принимая, что все $p_k > 0$).

Процедура моделирования сводится к моделированию дискретного распределения $\{p_k\}$ — выбору случайного номера k_0 , а затем к моделированию $\varphi_{k_0}(x)$. $\square$

Если φ обладает некоторыми свойствами гладкости ($\varphi \in \mathbf{C}^m$), то она может быть приближена ступенчатой функцией (рис. 2.3), так что площадь “между” графиком этой ступенчатой функции и кривой $y = \varphi(x)$ будет малой ($\sim O(N^{-m})$).

Тогда мы можем свести задачу к моделированию ступенчатой плотности с вероятностью $\sum_{k=1}^N |s_0|$ и с вероятностью $O(N^{-m})$ плотности, получаемой после нормировки заштрихованной части на

рис. 2.3. Здесь s_k обозначает площадь прямоугольника s_k . Моделирование ступенчатой плотности, в свою очередь, сводится к моделированию дискретного распределения и требует $\log_2 N$ операций. При достаточно большом N это и есть оценка трудоемкости алгоритма. Очевидно, возможны и другие виды аппроксимации разбиений плотностей. Аналогичные подходы возможны и в многомерном случае, хотя строить ступенчатые аппроксимации здесь значительно сложнее.

2.1.3. Метод Уокера

Как уже отмечалось, предварительная обработка распределений может служить средством, резко повышающим эффективность моделирования. В связи с этим значительный интерес представляет метод, предложенный Уокером [27]. Он использует то обстоятельство, что моделирование дискретного распределения с равными вероятностями исходов осуществляется особенно просто. Действительно, если $p_i = 1/n$, $i = 1, \dots, n$, то реализация соответствующей целочисленной случайной величины ξ может быть получена по формуле $\xi = \lfloor n\alpha \rfloor + 1$.

Пусть теперь $p_1, \dots, p_n$ — произвольное распределение случайной величины, принимающей значения $1, \dots, n$ (n конечно). Справедливо следующее утверждение.

Теорема 2.2. Любое дискретное распределение с конечным числом n исходов может быть представлено в виде смеси равновероятных распределений с двумя исходами.

Доказательство. При $n = 1$ утверждение очевидно.

Задано распределение

$$p_1, \dots, p_n, n > 1, p_i > 0, \sum_{i=1}^n p_i = 1.$$

Отметим прежде всего:

- существует $i = i_0$ такое, что $p_{i_0} \leq 1/n$. В противном случае имеем $\sum_{i=1}^n p_i > \sum_{i=1}^n 1/n = 1$.
- существует $i = j_0$ такое, что $p_{i_0} + p_{j_0} > 1/n$. Если это не так, то $p_i + p_j \leq 1/n$ для всех $i \neq j$ и $\sum_{i \neq j} (p_i + p_j) \leq \frac{n-1}{n} < 1$, но $\sum_{i \neq j} (p_i + p_j) = (n-1)p_i + 1 - p_i \geq 1$.

Теперь будем преобразовывать исходное распределение следующим образом. Введем верхний индекс, полагая $p_l = p_l^{(n)}$, $l = 1, \dots, n$.

Для $n \geq 2$ выберем i и $j \neq i$ так, чтобы $p_i^{(n)} \leq 1/n$ и $p_i^{(n)} + p_j^{(n)} > 1/n$, и положим

$$\begin{aligned} q_i^{(1)} &= np_i^{(n)}, & p_i^{(n-1)} &= 0, \\ q_j^{(1)} &= 1 - np_i^{(n)}, & p_j^{(n-1)} &= \frac{n}{n-1} \left(p_i^{(n)} + p_j^{(n)} - \frac{1}{n} \right), \\ q_k^{(1)} &= 0, & p_k^{(n-1)} &= \frac{n}{n-1} p_k^{(n)}, \quad k \neq i, k \neq j. \end{aligned} \quad (2.1.13)$$

Тогда для всех $l = 1, \dots, n$ имеем

$$p_l^{(n)} = \frac{1}{n} + \frac{n-1}{n} p_l^{(n-1)}.$$

Далее определим числа $q_i^{(2)}$ и $p_j^{(n-2)}$ по формулам (2.1.13), заменяя в них n на $n-1$ так, что

$$p_l^{(n-1)} = \frac{1}{n-1} q_l^{(2)} + \frac{n-2}{n-1} p_l^{(n-2)}.$$

Повторяя процедуру n раз:

$$p_l^{(n-j)} = \frac{1}{n-j} q_l^{(j+1)} + \frac{n-j-1}{n-j} p_l^{(n-j)},$$

получим

$$p_l^{(n)} = \frac{1}{n} \sum_{j=1}^n q_l^{(j)},$$

где $q_l^{(j)}$ – дискретное распределение с двумя исходами. $\square$

2.1.4. Битовые алгоритмы для моделирования дискретных распределений

Если имеется последовательность случайных (псевдослучайных) битов, порождаемых, например, алгоритмом (1.3.9), и имеется алгоритм, преобразующий их в биты реализации случайной величины с заданным законом распределения, то такого рода алгоритм мы будем именовать битовым. В качестве некоторой абстрактной модели можно говорить о бросании монеты и преобразовании последовательности нулей и единиц, соответствующих исходам бросаний, в последовательность битов нужной нам реализации случайной величины. Битовые алгоритмы изучались Кнутом и Яо [54]. Простейший пример, приведенный в их работе, — это имитация бросания игральной кости с помощью бросания монеты. Бросание идеальной кости есть механизм,

позволяющий получить с равной вероятностью один из 6 исходов. Моделируется дискретное распределение

1	2	3	4	5	6
1/6	1/6	1/6	1/6	1/6	1/6

Троекратное бросание монеты (1 — орел, 0 — решка) дает такое распределение вероятностей

000	001	010	011	100	101	110	111
1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

(2.1.14)

Дополним исходное распределение до распределения

0	1	2	3	4	5	6	7
0	1/6	1/6	1/6	1/6	1/6	1/6	0

(2.1.15)

Видим, что распределение (2.1.15) абсолютно непрерывно по отношению к распределению (2.1.14). Здесь можно использовать метод отбора. Его эффективность есть $3/4$, а среднее число бросаний монеты до получения результата равно

$$\frac{3}{\text{эффективность}} = \frac{3}{3/4} = 4.$$

С другой стороны, исходы, соответствующие 7(111) и 0(000), равновероятны и могут не отбрасываться, а служить началом нового разветвления (то есть мы уже имеем первую двоичную цифру новой тройки). В этом случае среднее число бросаний уменьшается до $3\frac{2}{3}$. Действительно, бросив первый раз монету трижды, мы затем будем бросать ее лишь дважды. Это и дает $3 \cdot 1 + 2(\frac{1}{3/4} - 1) = 3\frac{2}{3}$. Процедуру моделирования можно представить в виде бинарного дерева.

Если пользоваться описанным экономичным способом, требующим в среднем $3\frac{2}{3}$ бросания монеты (стрелочки возврата на рис. 2.4), то дерево может быть развернуто в бесконечное дерево с периодической структурой.

Легко показать (см. [54]), что любому дискретному распределению можно сопоставить, и не единственным образом, такое бинарное

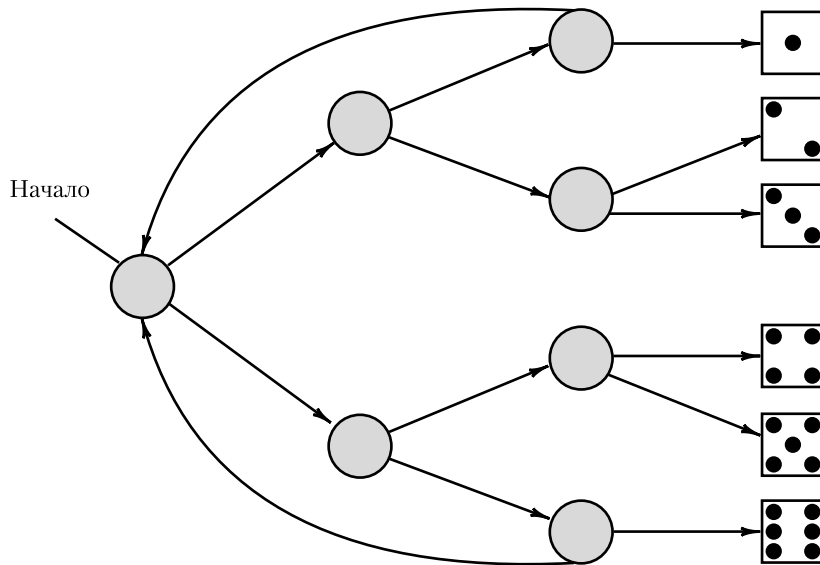


Рис. 2.4

дерево, называемое ПДР-деревом (порождающим дискретное распределение). В общем случае дерево бесконечно и не обязательно имеет периодическую структуру. Дерево, очевидно, однозначно связано с алгоритмом моделирования (ПДР-алгоритмом).

Оптимальным в смысле Кнута и Яо алгоритмом моделирования называется ПДР-алгоритм, минимизирующий среднее число (время) бросания монеты (если оно конечно).

Задача построения оптимального ПДР-алгоритма для дискретных распределений решена в цитированной выше работе. Дерево (алгоритм) рисунка является оптимальным. Изложение результатов работы Кнута и Яо потребовало бы много времени, и мы ограничимся лишь формулировкой основных понятий, связанных с оптимизацией битовых алгоритмов.

Остановимся более подробно на схеме бросания монеты (она же — хорошо исследованная в теории вероятностей схема Бернулли с равновероятными исходами). Повторение эксперимента порождает двоичное дерево с 2^M равновероятными исходами (концевые узлы) при M последовательных независимых бросаниях. Пусть теперь задано распределение вероятностей $\pi = \{p_i\}_{i=1}^N$ событий, занумерованных

числами $1, \dots, N$, и p_i являются двоично рациональными дробями, имеющими представление

$$p_i = 0_2 \pi_1^i \dots \pi_m^i, \quad i = 1, 2, \dots, N.$$

Если теперь при бросании монеты мы разделим концевые узлы на непересекающиеся подмножества A_i так, что в A_i войдет

$$\pi_1^i 2^{m-1} + \pi_2^i 2^{m-2} + \dots + \pi_m^i \quad (2.1.16)$$

концевых узлов, то, регистрируя попадание концевого узла в A_i , получим модель распределения π . (Считаем, что произошло событие с номером i .)

В частности, можно организовать отбор поддеревьев основного дерева с концевыми узлами в A_i следующим образом.

Если при заданном k ($k \leq M$) для некоторых значений $i = i_1, \dots, i_m$ ($m < 2^k$) оказывается $\pi_k^{i_l} = 1$, $l = 1, \dots, m$, то мы помечаем m узлов k -го уровня и считаем, что все “потомки” M -го поколения l -го узла принадлежат множеству A_{i_l} . Таким образом, можно прекращать просмотр (реализацию) двоичного дерева по достижении помеченного узла. Легко видеть, что p_i соответствует ровно (2.1.16) из 2^M исходов и эта процедура выделяет ПДР-дерево, содержащее минимальное число концевых узлов.

Одним из важных результатов, полученных в рамках битового подхода, является следующий.

Если $p_1, \dots, p_n$ — дискретное распределение, то среднее число битов, необходимых для получения реализации этого распределения $\nu(p_1, \dots, p_n)$, не превосходит величины $\log_2(n-1) + 2$ (см. [54], с. 113).

Мы ограничились изложением наиболее употребительных общих методов моделирования распределений. Под общими мы понимаем методы, применимые к любым распределениям или, по крайней мере, к очень широким классам распределений.

В начале главы 1 были сформулированы свойства равномерно распределенных случайных величин, позволяющие, в частности, оценивать интегралы высокой кратности от интегрируемых по Риману функций. Моделирование случайных величин с заданным законом распределения позволяет с помощью суммы $\frac{1}{N} \sum_{j=1}^N f(Y_j)$, где Y_j — независимые реализации случайного вектора с плотностью $p(Y)$, оценивать интеграл $\int f(X)p(x) dX$. Обоснованием такой оценки очевидным образом служит также закон больших чисел.

Однако статистическое моделирование (метод Монте-Карло) наиболее эффективно при решении еще более сложных задач — решении больших систем уравнений, моделировании сложных систем.

Используя изложенные выше методы, мы можем создавать имитационные модели явлений, допускающих теоретико-вероятностное описание. Некоторые простые примеры приводятся ниже. Подробнее см., например, [35].

2.2. Моделирование случайных процессов. Интегралы по траекториям

Случайные процессы — тема весьма обширная. Мы рассмотрим исключительно алгоритмические аспекты, избегая общего абстрактного подхода. В большинстве случаев случайный процесс будет описан как алгоритм моделирования. Конечно, мы не сможем избежать упоминания некоторых известных фактов математической теории, но не станем их доказывать (см., например, [31]).

Приведем некоторые основные определения.

Определение 2.1. Случайным процессом на вероятностном пространстве $(\Omega, \mathcal{A}, P)$ называется семейство случайных величин $\xi(t, w)$, зависящих от вещественного параметра t , $t \in \mathbf{R}^1$.

Определение 2.2. Пусть $t_1, t_2, \dots, t_n$ — некоторые значения параметра t из множества H . Совместные распределения случайных величин $\xi(t_1, w), \xi(t_2, w), \dots, \xi(t_n, w)$ при всевозможных n и наборах $t_1, \dots, t_n$ называются конечномерными распределениями вероятностей случайного процесса.

Согласно теореме Колмогорова, заданное для всевозможных n и $t_1, \dots, t_n$ семейство функций распределения $F_n(x_1, t_1, \dots, x_n, t_n) \rightarrow P(\xi(t_1, w) \leq x_1, \dots, \xi(t_n, w) \leq x_n)$ определяет случайный процесс (то есть существует вероятностное пространство $(\Omega, \mathcal{A}, P)$) и семейство измеримых функций $\xi(t, w), w \in \Omega$ (случайный процесс), если выполнимы условия согласования:

$$\lim_{x_n \rightarrow \infty} F_n(x_1, t_1, \dots, x_n, t_n) = F_{n-1}(x_1, t_1, \dots, x_{n-1}, t_{n-1}). \quad (2.2.1)$$

Таким образом, задача моделирования случайного процесса может быть сведена к задаче моделирования заданного семейства распределений.

В большинстве случаев, представляющих прикладной интерес, существуют конечномерные плотности совместного распределения для

случайного процесса, и мы имеем задачу моделирования этих плотностей.

Пусть, например, H есть отрезок $[0, T]$ вещественной прямой. Если для некоторого процесса $\xi(t, w)$ заданы всевозможные конечномерные плотности распределения

$$\varphi(x_1, t_1, \dots, x_n, t_n), \quad (2.2.2)$$

то, фиксируя $t_1, \dots, t_n$ и моделируя полученную плотность совместного распределения случайных величин $\xi(t_1, w), \dots, \xi(t_n, w)$ методами предыдущего раздела, мы получим реализацию этих случайных величин. Выбирая, например $t_j = \frac{1}{n} \cdot j$, $j = 0, \dots, n-1$, можно при достаточно большом n проследить подробно поведение одной из реализаций случайного процесса. Проблема, однако, состоит в том, что моделирование случайного вектора большой размерности, как мы видели, может быть очень трудной (трудоемкой) задачей. Иногда эти трудности непреодолимы даже с помощью современной вычислительной техники.

На практике обычно имеют дело с некоторыми сравнительно узкими классами случайных процессов, для которых либо плотность (2.2.2) распадается на произведение плотностей, зависящих от небольшого числа переменных, либо существуют простые методы моделирования этой плотности.

Для того чтобы выделить удобные для моделирования процессы, напомним следующие определения.

Определение 2.3. Функция параметра t , равная величине $E\xi(t, w)$ при каждом фиксированном t , называется математическим ожиданием или средним значением случайного процесса, а функция двух переменных $R(\theta, t) = E(\xi(\theta, w)\xi(t, w))$ — ковариационной функцией.

Ниже даются определения важных частных случаев случайных процессов. Процессы определяются в узком смысле (см. [31]).

Определение 2.4. Случайный процесс $\xi(\theta) = \xi(\theta, w)$ называется стационарным, если связанные с ним распределения вероятностей не зависят от сдвига параметра θ . Если все конечномерные распределения процесса нормальны, то процесс называется гауссовским случайным процессом.

Определение 2.5. Марковским случайным процессом называется процесс, для условных конечномерных распределений

которого с вероятностью 1 выполняются равенства

$$P\{\xi(\theta_n) \leq x : \xi(\theta_1), \dots, \xi(\theta_{n-1})\} = P\{\xi(\theta_n) \leq x : \xi(\theta_{n-1})\}$$

для любых значений параметра $\theta_1 \leq \dots \leq \theta_n$ из Θ . Это условие эквивалентно следующему: для любых $t_1 < t_2$ из Θ и любого x с вероятностью 1

$$P\{\xi(t_2) \leq x : \xi(t), t \leq t_1\} = P\{\xi(t_2) \leq x : \xi(t_1)\}.$$

Марковский случайный процесс может быть задан переходными вероятностями и начальным распределением вероятностей. В этом случае предполагается, что Θ имеет минимальный элемент Θ_0 , при котором заданы начальное распределение вероятностей p_0 случайной величины $\xi(\theta_0)$ и функция (переходная вероятность) $P(w, t; B, \theta)$, $w, w' \in \Omega$; $t, \theta \in \Theta$; $B \in \mathcal{A}$, удовлетворяющая для почти всех w равенству

$$P(w, t; B, \theta) = \int P(w', \tau; B, \theta) P(w, t; dw', \tau) \quad (2.2.3)$$

(уравнение Чепмена–Колмогорова). Последнее обеспечивает выполнение условий согласованности конечномерных распределений процесса.

Мы ограничимся рассмотрением только отдельных частных случаев упомянутых двух типов случайных процессов. Это позволит выяснить некоторые особенности возникающих алгоритмов моделирования и обсудить важные приложения построенных моделей.

Сначала обсудим алгоритмическую модель марковского процесса с дискретным временем. Не умаляя общности, положим, что параметр t принимает значения $t = 0, 1, 2, \dots$. Пусть $\mathcal{X}$ есть фазовое пространство значений марковского процесса $\xi(t)$. Однородный процесс в соответствии со сказанным выше задается начальным распределением своих значений при $t = 0$ и при каждом $t > 0$ — условным распределением значений $\xi(t+1)$ при условии, что $\xi(t)$ имеют заданное значение. Процесс может задаваться соответствующими плотностями распределения $p^0(x)$ и $p(x, y)$ по отношению к мере Лебега^{*)}.

Легко видеть, что реализация процесса, или его траектория — последовательность величин $x_0, x_1, x_2, \dots$, вычисляется следующим образом. Методами п. 2.1 моделируем x_0 , распределенный с плотностью $p^0(x)$. Затем если имеем некоторый x_t , то x_{t+1} получается в результате моделирования плотности $p(x_t, x_{t+1})$ при фиксированном x_t .

^{*)} Рассматривается случай, когда p не зависит от времени.

Таким образом можно получить, вообще говоря, сколь угодно длинную последовательность x_t . В некоторых случаях, однако, существует такое $t = \tau$, что оказывается: $x_\tau = x_{\tau+1} = x_{\tau+2} = \dots$, при этом говорят, что x_τ принадлежит множеству поглощающих состояний a . Говорят также, что траектория процесса обрывается при попадании в множество a . Если вероятность попадания в это множество положительна, то плотность перехода нормирована следующим образом:

$$\int_{\mathcal{X}} p(x, y) dy = 1 - g(x), \quad (2.2.4)$$

где $g(x)$ — вероятность поглощения (обрыва траектории) и $0 \leq g(x) \leq 1$. Если $g(x) > 0$ для некоторых x , то моделирование процесса осуществляется следующим образом.

При $t = 0$, x_0 получим, моделируя плотность $p^0(x)$. Если получено x_t , то с вероятностью $g(x_t)$ траектория процесса обрывается, и с вероятностью $1 - g(x_t)$ находим x_{t+1} , распределенный с плотностью

$$\frac{p(x_t, x_{t+1})}{1 - g(x_t)}, \quad (2.2.5)$$

x_t — фиксирован.

Если множество значений процесса $\xi(t)$ конечно или счетно, то соответствующие распределения дискретны. Процесс — цепь Маркова с конечным или счетным числом состояний — задается вектором начального распределения $p^0 = (p_1^0, \dots, p_n^0)$ и матрицей перехода $\mathcal{P} = \|p_{i,j}\|_{i,j=1}^n$, $p_{i,j} \geq 0$:

$$\sum_{j=1}^n p_{i,j} = 1 - g_i, \quad 0 \leq g_i \leq 1, \quad (2.2.6)$$

где g_i — вероятность обрыва траектории, если цепь находится в состоянии с номером i ; а $p_{i,j}$ — вероятность перехода в состояние с номером j из состояния с номером i (условная вероятность перехода в состояние j в момент времени $t+1$, если в момент времени t цепь находилась в состоянии с номером i).

Таким образом при $t = 0$ моделируется распределение p^0 . В результате моделирования мы получаем номер состояния i_0 . Далее, если в момент времени t цепь находится в состоянии i_t , то с вероятностью g_{i_t} траектория обрывается и с вероятностью $1 - g_{i_t}$ находится следующий номер состояния i_{t+1} , распределение которого есть

$$\frac{1}{1 - g_{i_t}}(p_{i_t,1}, \dots, p_{i_t,n}). \quad (2.2.7)$$

Результатом моделирования является последовательность номеров

$$i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_t \rightarrow \dots \quad (2.2.8)$$

— траектория процесса. Условия, при которых все траектории марковского процесса обрываются с вероятностью 1, зависят от структуры матрицы $\mathcal{P}$, а в непрерывном случае — от свойств переходной вероятности $p(x, y)$.

Из теории марковских процессов следует (в частности, из (2.2.3)), что распределение вероятностей состояний $p_i(t)$ в момент времени t подчиняется уравнению

$$p_i(t+1) = \sum_j p_{j,i} p_j(t) + g_i \chi_a \quad (2.2.9)$$

или

$$p(t+1, x) = \int_{\mathcal{X}} p(x, y) p(t, y) dy + g(x) \chi_a \quad (2.2.10)$$

— в непрерывном случае. Здесь $p_i(t) = \mathbf{P}\{i_t = i\}$, $p(t, y) = \mathbf{P}(x_t = y)$, а χ_a — характеристическая функция поглощающего состояния. Анализируя уравнения (2.2.9) и (2.2.10), легко усмотреть, что по крайней мере при $g_i > 0$, $i = 1, \dots, n$, а также при $g(x) > 0$, $x \in \mathcal{X}$, почти все траектории процесса попадают в поглощающее состояние.

Рассмотрим теперь гауссовский случайный процесс $\xi(t)$. Пусть на промежутке $[0, T]$ заданы моменты времени $0 < t_1 < t_2 < \dots < t_{n-1} < T$; $\xi(t)$ полностью определяется заданием своих первых двух моментов: математическим ожиданием

$$\mathbf{E}\xi(t) = f(t) \quad (2.2.11)$$

и ковариационной функцией

$$\mathbf{E}[(\xi(t) - f(t))(\xi(\tau) - f(\tau))] = \mathbf{R}(t, \tau). \quad (2.2.12)$$

Распределение случайного вектора $\Xi = (\xi(0), \xi(t_1), \dots, \xi(t_{n-1}), \xi(T))$ есть многомерное нормальное распределение, задаваемое плотностью

$$\varphi_{n+1}(X) = (2\pi)^{-(n+1)/2} (\det C)^{-1/2} \exp\left(-\frac{1}{2}(X - F)^T C^{-1}(X - F)\right), \quad (2.2.13)$$

где $X = (x_0, \dots, x_n)^T$, $F = (f(0), f(t_1), \dots, f(T))^T$, $C = \|\mathbf{R}(t_i, t_j)\|_{i,j=0}^n$, $t_0 = 0$, $t_n = T$, $\det C \neq 0$.

Таким образом, задача моделирования гауссовского процесса в выбранных точках t_i , $i = 0, \dots, n$, сводится к моделированию многомерной плотности (2.2.13). Алгоритм состоит из подготовительной стадии (разложение матрицы C на произведение двух треугольных) и основной (вычисление линейной комбинации $n+1$ независимых случайных нормально $(0, 1)$ распределенных величин).

Матрица C положительно определена и симметрична и может быть представлена в виде

$$C = \Gamma \Gamma^T, \quad (2.2.14)$$

где Γ — верхняя треугольная матрица:

$$\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1,n+1} \\ 0 & \gamma_{22} & \dots & \gamma_{2,n+1} \\ 0 & 0 & & \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \gamma_{n+1,n+1} \end{pmatrix}$$

Величины $\gamma_{i,j}$ вычисляются с использованием “метода квадратных корней”, хорошо известного в линейной алгебре [68]. Этот алгоритм легко построить, если заметить, что числа $\gamma_{i,j}$ при заданных $c_{i,j}$ определяются последовательно из цепочки уравнений, получаемой приравнованием соответствующих элементов матриц в правой и левой частях равенств (2.2.14).

Обозначим далее $X - F = Y$. Имеем $(X - F)^T C^{-1} (X - F) = Y^T C^{-1} Y = Y^T (\Gamma \Gamma^T)^{-1} Y = Y^T (\Gamma^{-1})^T \Gamma^{-1} Y = (\Gamma^{-1} Y)^T (\Gamma^{-1} Y)$.

Если теперь в (2.2.13) сделать замену переменных $Z = (\Gamma^{-1} Y)$, то получим

$$\tilde{\varphi}_{n+1}(Z) = (2\pi)^{-(n+1)/2} \exp\left(-\frac{1}{2} Z^T Z\right).$$

Случайный вектор с плотностью распределения $\tilde{\varphi}_{n+1}(Z)$ является $(n+1)$ -мерным нормальным случайным вектором с независимыми компонентами $\xi = (\xi_1, \dots, \xi_{n+1})$, $\xi_i \in N(0, 1)$. Моделирование ξ_i может осуществляться методом, описанным в примере на с. 38. Затем вектор $\Xi = [\xi(0), \xi(t_1), \dots, \xi(t_{n-1}), \xi(T)]$ находят по формуле

$$\Xi = F + \Gamma \xi. \quad (2.2.15)$$

Пусть $t_i = T \frac{i}{n}$. При этом условии матрица ковариаций обычно имеет приближенно ленточную структуру: зависимость $\xi(t_i)$ от “далеких” значений убывает с их удаленностью. В этом случае можно

построить удобные специальные алгоритмы (авторегрессионные модели и т. п.).

Примером гауссовского марковского процесса с непрерывным временем может служить процесс одномерного броуновского движения. Если частица начинает движение в момент времени $t = 0$ в точке $x = 0$, то вероятность того, что в моменты времени $t_1, t_2, \dots, t_n$ ($0 < t_1 < \dots < t_n$) она будет находиться в промежутках $[a_1, b_1], \dots, [a_n, b_n]$ соответственно, равна

$$\int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} p(0, x_1; t) p(x_1, x_2; t_2 - t_1) \dots p(x_{n-1}, x_n; t_n - t_{n-1}) dx_1 \dots dx_n, \quad (2.2.16)$$

где

$$p(x, y; t) = (2\pi t)^{-1/2} \exp\left[-\frac{1}{2t}(y - x)^2\right], \quad (2.2.17)$$

что и определяет конечномерные распределения процесса броуновского движения. Благодаря марковости и специальному виду $p(x, y; t)$ моделирование движения броуновской частицы может осуществляться путем последовательного увеличения числа точек, в которых моделируется процесс.

Пусть $\xi(t)$ — траектория броуновской частицы и $0 \leq t \leq 1$, $\xi(0) = 0$. Будем приближать $\xi(t)$ ломаными, вершины которых поместим в точках $t_i = i/2^n$, $i = 0, 1, \dots, 2^n$. Заметим прежде всего, что при $n = 1$ случайная величина $\xi(1)$ распределена нормально со средним 0 ($\xi(0) = 0$) и дисперсией 1, и ее реализация может быть получена одним из обсуждавшихся выше способов. Увеличение числа n не влечет за собой пересчета полученных ранее реализаций. При $n = 2$ случайная величина $\xi(1/2)$ распределена нормально со средним $(\xi(0) + \xi(1))/2$ и дисперсией $1/4$. Вообще, если получены реализации $\xi(t')$ и $\xi(t'')$, $t'' > t'$, то $\xi(\frac{t' + t''}{2})$ имеет нормальное распределение со средним $(\xi(t') + \xi(t''))/2$ и дисперсией $(t'' - t')/2$. Это позволяет последовательно дробить промежуток $[0, 1]$ на части, кратные 2^n , строить ломаные с соответствующим числом звеньев и следить за сходимостью интересующего нас функционала на траекториях к своему предельному значению.

Существует много приемов моделирования специальных классов случайных процессов, однако мы предполагали рассмотреть только простейшие из них.

Как правило, моделирование процесса не является самоцелью. Вычисляется функционал — число, сопоставляемое данной траектории процесса. Задачи такого рода могут иметь большое прикладное

значение. В качестве примера снова рассмотрим однородную цепь Маркова с конечным числом n состояний. Считаем, что время дискретно и траектории цепи обрываются с вероятностью 1.

Пусть цепь задается, как и ранее, начальным распределением p^0 и переходной матрицей $\mathcal{P} = \|p_{i,k}\|$; $\sum_{k=1}^n p_{i,k} = 1 - g_i$. Для простоты изложения полагаем $g_i > 0$, $i = 1, \dots, n$. Полученной в результате моделирования траектории

$$i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_\tau \quad (2.2.18)$$

соответствует вероятность

$$p_{i_0}^0 p_{i_0, i_1} \dots p_{i_{\tau-1}, i_\tau} g_{i_\tau}. \quad (2.2.19)$$

Таким образом на множестве траекторий определяется вероятностная мера. Действительно, имеет место равенство

$$\sum_{\tau=0}^{\infty} \sum_{i_0=1}^n \dots \sum_{i_\tau=1}^n p_{i_0}^0 p_{i_0, i_1} \dots p_{i_{\tau-1}, i_\tau} g_{i_\tau} = 1. \quad (2.2.20)$$

Чтобы доказать это равенство, заметим, что система линейных алгебраических уравнений

$$Z = \mathcal{P}Z + G, \quad G = \|g_1, \dots, g_n\|^T$$

имеет единственное итерационное решение. Это так, ибо первая норма матрицы $\mathcal{P}$, равная $\max_i \sum_{k=1}^n p_{i,k}$, меньше единицы ($g_i > 0$). Легко видеть, с одной стороны, что это решение есть вектор, все компоненты которого равны 1. Действительно, имеем

$$1 = \sum_{k=1}^n p_{i,k} \cdot 1 + g_i, \quad i = 1, \dots, n.$$

Равенство выполняется в силу определения g_i . С другой стороны, решение системы есть $Z = \lim_{m \rightarrow \infty} Z_m$, где $Z_m = \mathcal{P}Z_{m-1} + G$, $Z_0 = G$. Но $Z_m = (I + \mathcal{P} + \dots + \mathcal{P}^m)G$ и легко проверить, что

$$Z = \left\| \sum_{\tau=1}^{\infty} \sum_{i_1=1}^n \dots \sum_{i_\tau=1}^n p_{i_0, i_1} \dots p_{i_{\tau-1}, i_\tau} g_{i_\tau} \right\|_{i_0=1}^n, \quad (2.2.21)$$

откуда, в свою очередь, следует $(p^0, Z) = 1$, что совпадает с (2.2.20), если учесть (2.2.21). Таким образом, множество траекторий, с определенной на нем мерой μ согласно (2.2.18), (2.2.19), есть вероятностное пространство.

Если на траекториях (2.2.18) задан функционал $\Phi(x_0, x_1, \dots, x_\tau)$ и существует его математическое ожидание, то оно выражается формулой

$$E\Phi = \sum_{\tau=0}^{\infty} \sum_{i_0=1}^n \dots \sum_{i_\tau=1}^n \Phi(x_0, \dots, x_i) p_{i_0}^0 p_{i_0, i_1} \dots p_{i_{\tau-1}, i_\tau} g_{i_\tau}. \quad (2.2.22)$$

Один из широко известных фактов, связанных с представлением (2.2.22), относится к вероятностным методам решения систем линейных алгебраических уравнений.

Пусть задана система линейных уравнений вида

$$X = AX + F, \quad (2.2.23)$$

где $A = \|a_{ij}\|_{i,j=1}^n$, $X = (x_1, \dots, x_n)^T$, $F = (f_1, \dots, f_n)^T$, причем сходится метод последовательных приближений

$$X_m = AX_{m-1} + F, \quad X_0 = F, \quad m = 1, 2, \dots$$

Это означает, что итерационное решение $\tilde{X}$ системы (2.2.23) существует, единственно и может быть представлено в виде

$$\tilde{X} = \sum_{k=0}^{\infty} A^k F \quad (\text{сумма ряда Неймана}).$$

Если выполнены более жесткие условия — условия сходимости так называемого мажорантного итерационного процесса:

$$\tilde{X}_m = |A| \tilde{X}_{m-1} + |F|, \quad |A| = \left\| |a_{ij}| \right\|_{i,j=1}^m,$$

то $\tilde{X}$ представимо в форме

$$\tilde{X} = \left\| \sum_{k=0}^{\infty} \sum_{i_1=1}^n \dots \sum_{i_k=1}^n a_{i,i_1} a_{i_1,i_2} \dots a_{i_{k-1},i_k} f_{i_k} \right\|_{i=1}^n,$$

а скалярное произведение $(H, \tilde{X})$, где $H = (h_1, \dots, h_n)$ — произвольный заданный вектор в форме

$$(H, \tilde{X}) = \sum_{k=0}^{\infty} \sum_{i_0=1}^n \dots \sum_{i_k=1}^n h_{i_0} a_{i_0,i_1} \dots a_{i_{k-1},i_k} f_{i_k}. \quad (2.2.24)$$

Достаточно сравнить (2.2.22) с (2.2.24), чтобы указать вероятностный алгоритм вычисления $(H, \tilde{X})$.

Действительно, пусть $p^0, \mathcal{P}$ — произвольная цепь Маркова, удовлетворяющая условиям согласования:

$$\begin{aligned} p_i^0 &> 0, & \text{если } h_i &\neq 0, & i &= 1, \dots, n, \\ p_{i,j} &> 0, & \text{если } a_{i,j} &\neq 0, & i, j &= 1, \dots, n, \\ g_i &> 0, & \text{если } f_i &\neq 0, & i &= 1, \dots, n. \end{aligned} \quad (2.2.25)$$

Тогда, выбирая, например,

$$\Phi(x_0, \dots, x_\tau) = \Phi_0(x_0, \dots, x_\tau) = \frac{h_{i_0} a_{i_0, i_1} \dots a_{i_{\tau-1}, \tau} f_{i_\tau}}{p_{i_0}^0 p_{i_0, i_1} \dots p_{i_{\tau-1}, \tau} g_{i_\tau}},$$

видим, что выполняется равенство

$$E\Phi_0 = (H, \tilde{X}), \quad (2.2.26)$$

причем существование математического ожидания следует из сходимости мажорантного итерационного процесса. Моделируя N независимых траекторий цепи Маркова и вычисляя для каждой значение Φ_0^j (j — номер траектории, $j = 1, \dots, N$), можно на основании закона больших чисел оценить $(H, \tilde{X})$ с помощью среднего арифметического

$$\frac{1}{N} \sum_{j=1}^N \Phi_0^j.$$

Таким образом, мы видим, что вероятностные методы вычислений пригодны не только для оценки интегралов высокой кратности, но и для вычисления существенно более сложных объектов — интегралов по траекториям случайных процессов.

2.3. Библиографическое дополнение к главе 2

В главе содержатся элементарные сведения о моделировании распределений с заданной плотностью. Описаны общие (то есть относящиеся к широкому классу распределений) методы, которые позволяют построить в каждом конкретном случае достаточно эффективную процедуру моделирования. Большинство приемов повышения эффективности излагается с единой точки зрения “повышения размерности”.

Огромное количество конкретных приемов моделирования случайных величин и векторов с заданной плотностью содержит книга Девроя [4]. На русском языке процедуры моделирования ряда конкретных распределений и случайных процессов можно найти в учебнике Михайлова и Войтишека [59]. Многие методы моделирования распределений с подробным описанием алгоритмов представлены также в книге Фишмана [9]. Моделированию случайных процессов посвящена обширная монография Пригарина [63].

Выбор оптимального метода моделирования данного распределения до сих пор не формализован и является своего рода искусством. Исключение составляет подход, развитый Кнутом и Яо [20], есть и перевод на русский язык [54]. Суть его изложена применительно к дискретным распределениям в разделе “Битовые алгоритмы”. Для дискретных распределений мы имеем в своем роде окончательные результаты. В общем случае намечены лишь основные подходы.

Многие прикладные задачи сводятся к вычислению средних значений функционалов на траекториях случайных процессов.

Как правило, речь идет о марковских процессах или процессах, которые являются марковскими в некотором специально выбранном пространстве состояний. Метод Монте-Карло широко применяется, в частности, для решения задач переноса излучения [56], массового обслуживания [9], [41]. При переходе к изучению стохастических методов численного интегрирования важно подчеркнуть, что речь идет не просто о вычислении интегралов конечной кратности, но о более общих объектах — интегралах по σ -конечной мере. В нашем случае мы ограничились случаем вероятностной меры, что принципиально не влияет на решение проблемы.

Интегралы по траекториям случайных процессов (континуальные интегралы) широко представлены в литературе. Упомянем книгу Егорова и др. [34], где имеются многочисленные ссылки. Приведем широко известный пример интеграла по траекториям, являющимися непрерывными кривыми (интеграл Винера).

Пусть Ω — пространство непрерывных на $[0, \infty)$ функций $x(t)$ и $x_i = x(t_i)$, $i = 1, \dots, n$; $x(0) = 0$. Тогда формулы (2.2.16), (2.2.17) определяют вероятностную меру на полукольце множеств вида

$$[a_1 \leq x(t_1) \leq b_1; \dots; a_n \leq x(t_n) \leq b_n]$$

(выполнение условий согласованности легко проверить) и эта мера может быть распространена на все пространство Ω непрерывных функций. В соответствии с общей схемой введения интеграла можно далее определить интеграл, называемый интегралом Винера.

Для непрерывной неотрицательной функции $f(x)$, $x \in (-\infty, +\infty)$, интеграл Винера

$$J = \int \exp \left[- \int_0^t f((x(\tau))) d\tau \right] \mathbf{P}(dw),$$

где $\mathbf{P}$ — мера Винера, также можно рассматривать как предел

$$J = \lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp \left[-\frac{t}{n} \sum_{i=1}^n f(x_i) \right] p(0, x_1; t_1) \dots \\ \dots p(x_{n-1}, x_n; t_n - t_{n-1}) dx_1 \dots dx_n,$$

который с учетом (2.2.17) запишем в виде

$$J = \lim_{n \rightarrow \infty} (2\pi t n^{-1})^{-n/2} \int_{-\infty}^{+\infty} \dots \\ \dots \int_{-\infty}^{+\infty} \exp \left[-t n^{-1} \left\{ \sum_{i=1}^n f(x_i) + \frac{1}{2} \sum_{i=1}^n (x_i - x_{i-1})^2 / (t n^{-1}) \right\} \right] dx_1 \dots dx_n.$$

Это позволяет получить J в форме

$$\int \exp \left[- \int_0^t \left\{ \frac{1}{2} \left(\frac{dx}{d\tau} \right)^2 + f(x(\tau)) \right\} d\tau \right] d(\text{траектория}).$$

С мерами на пространствах функций связаны многие уравнения в частных производных и важные прикладные задачи. Вычисление интегралов по траекториям случайных процессов с помощью моделирования занимает заметное место в вычислительной математике.

Глава 3

Вычисление интегралов

3.1. Метод Монте-Карло

Существенным отличием проблем, возникающих в связи с использованием метода Монте-Карло, от близких по характеру изучаемых классической вычислительной математикой, является необходимость моделирования вероятностных распределений (мер). Как мы видели в предыдущей главе, меры, определенные на траекториях случайных процессов, бывают достаточно сложными. Обычно рассматривается задача их приближенного (как правило, в слабом смысле) моделирования. Может возникнуть необходимость определять меру и на объектах более общей природы (например, случайные поля). Далее ограничимся примерами задач, в которых проблема моделирования меры довольно просто решается средствами главы 2. Таким образом, речь будет идти о вычислении интегралов $J[f] = \int f(x)\mu(dx)$, где μ — вероятностная мера на некотором множестве $\mathfrak{X}$ элементов x . Предполагается, что $\mathfrak{X}$ снабжено некоторой σ -алгеброй и f есть μ -измеримая на $\mathfrak{X}$ функция.

Тогда если мы умеем достаточно точно вычислять независимые реализации $x_1, x_2, \dots$ случайной величины x , распределенной в соответствии с мерой μ , то с вероятностью 1 имеем

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N f(x_j) = \int f(x)\mu(dx) \quad (3.1.1)$$

— усиленный закон больших чисел (частный случай).

Если f интегрируема с квадратом ($f \in \mathbf{L}^2(\mu)$), то в качестве частного случая центральной предельной теоремы при тех же предположениях относительно x_j , что и выше, имеем

$$\lim_{N \rightarrow \infty} \mathbf{P}\left(\frac{\sqrt{N}}{\sigma_f} \left| \frac{1}{N} \sum_{j=1}^N f(x_j) - \int f(x)\mu(dx) \right| < y\right) = \sqrt{\frac{2}{\pi}} \int_0^y e^{-u^2} du, \quad (3.1.2)$$

где $\sigma_f^2 = \int f^2(x)\mu(dx) - \left(\int f(x)\mu(dx)\right)^2$, или с заданной вероятностью p неравенство

$$\left| \frac{1}{N} \sum_{j=1}^N f(x_j) - \int f(x)\mu(dx) \right| \leq \frac{y\sigma_f}{\sqrt{N}}, \quad (3.1.3)$$

где y определяется из условия $\int_0^y e^{-u^2} du = \sqrt{\frac{\pi}{2}}p$.

Таким образом, если интеграл $J[f]$ оценивается с помощью суммы $S_N[f] = \frac{1}{N} \sum_{j=1}^N f(x_j)$, то погрешность при заданном p убывает как $O(N^{-1/2})$.

Замечание. Если $f \in \mathbf{L}^p(\mu)$, $1 \leq p < 2$, то равенство (3.1.1) очевидным образом имеет место, но погрешность убывает медленнее чем $N^{-1/2}$. В частности, выполняется неравенство [1]

$$\mathbb{E}|S_N[f] - J|^p \leq \frac{2}{N^{p-1}} \sum_{k=1}^N \mathbb{E}|f(x_k) - J|^p. \quad (3.1.4)$$

□

Формулы (3.1.1) и (3.1.3) служат обоснованием использования суммы $S_N[f]$ в качестве приближенного значения интеграла J . Очень большое количество приложений использует подобную вычислительную схему.

Заметим сразу же, что $S_N[f]$ не является единственной несмещенной оценкой интеграла. Более того, она не является даже самой простой (в некотором разумном смысле) оценкой. Поясним это на простейшем примере интеграла $\int_0^1 f(x)dx$, где $f(x)$ — интегрируемая функция и $0 \leq f(x) \leq 1$ (рис. 3.1). Поскольку $\int_0^1 f(x)dx = \int_0^1 \int_0^{f(x)} dx dy$, то имеем

$$\mathbb{E}\left(\frac{1}{N} \sum_{i=1}^N f(x_i)\right) = \mathbb{E}\left(\frac{1}{N} \sum_{i=1}^N \zeta_i\right) = \int_0^1 f(x)dx,$$

где

$$\zeta_i = \begin{cases} 1, & \text{если } \alpha_{2i-1} \leq f(\alpha_{2i}), \\ 0, & \text{если } \alpha_{2i-1} > f(\alpha_{2i}), \end{cases}$$

$i = 1, \dots, N$ и $\alpha_1, \dots, \alpha_{2N}$ — независимые реализации равномерно распределенной на промежутке $[0, 1]$ случайной величины (в нашем случае псевдослучайные числа).

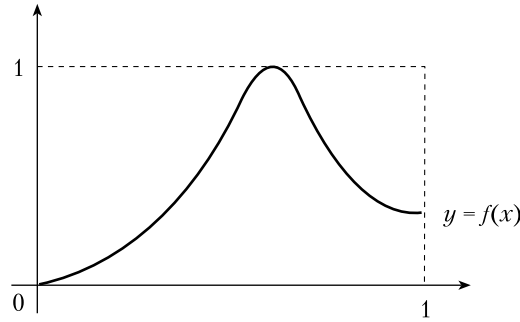


Рис. 3.1

При этом

$$D_1 = D\left(\frac{1}{N} \sum_{i=1}^N f(\alpha_i)\right) = \frac{1}{N} \left(\int_0^1 f^2(x) dx - \left(\int_0^1 f(x) dx \right)^2 \right),$$

$$D_2 = \frac{1}{N} (p(1-p)), \quad p = \int_0^1 f(x) dx$$

и, очевидно, $D_2 \geq D_1$ для всех f рассматриваемого нами вида.

Порядок убывания погрешности, определяемый формулой (3.1.3), считается очень медленным (уменьшение погрешности в 10 раз требует увеличения вычислительной работы в 100 раз). Эта трудность могла бы быть непреодолимой при решении многих задач, однако наличие множителя σ_f в правой части (3.1.3) указывает на возможность увеличения эффективности метода за счет построения оценок с возможно меньшей дисперсией, для которых вычисляемый интеграл является общим математическим ожиданием. Ниже рассмотрим некоторые распространенные методы понижения дисперсии оценок.

3.1.1. Метод выделения главной части

Очень простой метод, предполагающий, что заранее (аналитически или численно) определено значение интеграла $J_1 = \int g(x) \mu(dx)$, где g (как и f) интегрируема с квадратом по отношению к мере μ и g близка к f в метрике $\mathbf{L}^2(\mu)$. Тогда оценка

$$\varkappa_1 = J_1 + \frac{1}{N} \sum_{i=1}^N [f(x_i) - g(x_i)], \quad (3.1.5)$$

где x_i независимы и распределены в соответствии с мерой μ , является несмещенной оценкой интеграла $J[f]$. Ее дисперсия есть

$$D\mathfrak{X}_1 = \frac{1}{N} \left[\int (f - g)^2 \mu(dx) - (J[f - g])^2 \right]. \quad (3.1.6)$$

В соответствии с предположением о близости f к g эта дисперсия меньше, чем дисперсия $S_N[f]$. Такой простой прием иногда бывает очень полезным, но отметим сразу, что необходимость при каждом i вычислять значение $g(x_i)$ увеличивает вычислительную работу.

3.1.2. Замена меры (метод существенной выборки)

Рассмотрим один из наиболее употребительных методов понижения дисперсии. Предварительно следует напомнить определение производной Радона–Никодима.

Если σ -конечная мера μ допускает представление в виде

$$\mu(A) = \int_A f(x) \nu(dx),$$

$A \in Q$ (Q — σ -алгебра подмножеств рассматриваемого множества $\mathfrak{X}$), а f — неотрицательная интегрируемая функция, то $f(x)$ называют производной Радона–Никодима меры μ по мере ν и обозначают

$$f(x) = \frac{d\mu}{d\nu}(x).$$

Согласно теореме Радона–Никодима для заданных σ -конечных мер μ и ν производная $f(x)$ существует в том и только в том случае, если μ абсолютно непрерывна*) по отношению к ν ($\mu \ll \nu$), то есть из равенства $\nu(A) = 0$ следует $\mu(A) = 0$.

Пусть, например, вероятностная мера ν в единичном квадрате такова. С вероятностью p_1 мы имеем равномерное распределение в квадрате, с вероятностью p_2 — равномерное распределение на его диагонали ($x = y$), а с вероятностью p_3 — дискретное распределение в точках (x_i, y_i) , $i = 1, 2, 3, 4$, квадрата

Номер точки	1	2	3	4
Вероятность	0,1	0,1	0,2	0,6

$p_1 + p_2 + p_3 = 1.$

*) Понятие абсолютной непрерывности использовалось в п. 2.1.

Для того чтобы мера μ была абсолютно непрерывна по отношению к этой мере, необходимо, чтобы с вероятностью $q_1 > 0$ она имела плотность $\varphi_1(x, y)$, отличную от нуля в единичном квадрате, с вероятностью $q_2 > 0$ имела плотность $\varphi_2(x)$, отличную от нуля на диагонали, и с вероятностью $q_3 > 0$ имела в точках (x_i, y_i) распределение вида

Номер точки	1	2	3	4	, $q_{1,i} > 0, \sum_{i=1}^4 q_{1,i} = 1.$
Вероятность	q_{11}	q_{12}	q_{13}	q_{14}	

Тогда производная Радона–Никодима

$$\frac{d\mu}{d\nu}(x, y) = \begin{cases} \varphi_1(x, y) \cdot p_1/q_1, & x, y \in D, \\ \varphi_2(x, y) \cdot p_2/q_2, & x = y, \quad x, y \in D, \\ (p_3/q_3) \cdot (0, 1/q_{11}), & x = x_1, \quad y = y_1, \\ (p_3/q_3) \cdot (0, 1/q_{12}), & x = x_2, \quad y = y_2, \\ (p_3/q_3) \cdot (0, 2/q_{13}), & x = x_3, \quad y = y_3, \\ (p_3/q_3) \cdot (0, 6/q_{14}), & x = x_4, \quad y = y_4. \end{cases}$$

Таким образом, если ν — любая σ -конечная мера, по отношению к которой μ абсолютно непрерывна, то

$$J[f] = \int f(x) \mu(dx) = \int f(x) \frac{d\mu}{d\nu}(x) \nu(dx), \quad (3.1.7)$$

и для интеграла $J[f]$ может быть предложено множество несмещенных оценок

$$\varkappa_\nu[f] = \frac{1}{N} \sum_{j=1}^N f(Y_j) \frac{d\mu}{d\nu}(Y_j), \quad (3.1.8)$$

где Y_j — независимые случайные величины, распределенные в соответствии с мерой ν . Оценки $\varkappa_\nu[f]$ имеют общее математическое ожидание, не зависящее от ν , но легко видеть, что их дисперсии зависят от ν . Действительно

$$E\varkappa_\nu^2[f] = \frac{1}{N} \int f^2(x) \left(\frac{d\mu}{d\nu}(x) \right)^2 \nu(dx) = \frac{1}{N} \int f^2(x) \frac{d\mu}{d\nu}(x) \mu(dx). \quad (3.1.9)$$

Естественно, возникает задача: среди всех оценок вида (3.1.7) выбрать оценку с наименьшей дисперсией (выбрать меру ν , доставляющую оценке (3.1.7) минимальную дисперсию).

Для решения этой задачи отметим прежде всего, что имеет место неравенство Коши–Буняковского

$$\int \left(f(x) \frac{d\mu}{d\nu}(x) \right)^2 \nu(dx) \geq \left(\int |f(x) \frac{d\mu}{d\nu}(x)| \nu(dx) \right)^2 = \left(\int |f(x)| \mu(dx) \right)^2. \quad (3.1.10)$$

Таким образом, мы получили оценку снизу для $E\kappa_\nu^2[f]$.

Легко указать по крайней мере одну меру ν , определенную на множестве $B_f = \{x : f(x) \neq 0; \quad x \in \mathfrak{X}\}$, для которой в (3.1.10) имеет место равенство.

Это

$$\nu_0(dx) = \frac{1}{\bar{J}} |f(x)| \mu(dx), \quad x \in B_f, \quad \bar{J} = \int |f(x)| \mu(dx). \quad (3.1.11)$$

Действительно,

- (1) μ абсолютно непрерывна по отношению к ν_0 . Она отличается от ν_0 множителем $\bar{J}/|f(x)|$, $x \in B_f$;
- (2) имеет место равенство:

$$\int \left(f(x) \frac{d\mu}{d\nu}(x) \right)^2 \nu_0(dx) = \int \left(|f(x)| \frac{\bar{J}}{|f(x)|} \right)^2 \cdot \frac{1}{\bar{J}} |f(x)| \mu(dx) = \bar{J}^2,$$

откуда вытекает, что $E\kappa_\nu^2[f]$, а следовательно, и $D\kappa_\nu[f]$ при $\nu = \nu_0$ достигают нижней границы. Полученный результат позволяет сформулировать следующую теорему.

Теорема 3.1. (О существенной выборке [2].) Среди несмещенных оценок $J[f]$ вида

$$\kappa_\nu[f] = \frac{1}{N} \sum_{j=1}^N f(Y_j) \frac{d\mu}{d\nu}(Y_j),$$

где Y_j — независимые случайные величины, распределенные в соответствии с мерой $\nu(dx)$, оценка $\kappa_{\nu_0}[f]$, где

$$\nu_0(dx) = \frac{1}{\bar{J}} |f(x)| \mu(dx),$$

имеет наименьшую возможную дисперсию, равную

$$D\kappa_{\nu_0}[f] = \bar{J}^2 - J^2. \quad (3.1.12)$$

□

Следствие. Если $f(x) \geq 0$, $x \in \mathfrak{X}$, то имеем

$$D\kappa_{\nu_0}[f] = 0. \quad (3.1.13)$$

Последнее равенство очевидным образом вытекает из (3.1.12). $\square$

Обсуждение практических вопросов, связанных с выбором оптимальной плотности, проведем на простом случае, когда мера μ задается плотностью $\mu(dx) = \varphi(x)dx$, $x \in D_s \subset \mathbf{R}^s$.

Как в этом, так и в общем случае ясно, что моделирование случайной величины, имеющей распределение ν_0 , — задача столь же сложная, что и вычисление исходного интеграла.

Теорема 3.1, однако, может быть полезна в том отношении, что она рекомендует выбирать меру, близкую в некотором смысле к оптимальной. Пусть $\nu(dx) = \psi(x)\mu(dx)$. Тогда

$$E\kappa_\nu^2 = \int \left(f(x) \frac{\varphi(x)}{\psi(x)} \right)^2 \psi(x) \mu(dx)$$

и

$$\nu_0(dx) \equiv \psi_0(x)(dx) = \frac{1}{\bar{J}} |f(x)| \varphi(x) \mu(dx).$$

Будем рассматривать случай $f(x)\varphi(x) > 0$ и считать, что мы умеем моделировать плотность $\psi_\varepsilon(x) = \bar{J}_\varepsilon^{-1}(f(x)\varphi(x) + \varepsilon(x))$, где $\varepsilon(x)$ — малая добавка, $\bar{J}_\varepsilon = \int (f(x)\varphi(x) + \varepsilon(x))\mu(dx)$. Обозначая κ_ε оценку вида (3.1.8), соответствующую этой плотности, запишем

$$\begin{aligned} E\kappa_\varepsilon^2 &= \int \frac{f^2(x)\varphi^2(x)}{\psi_\varepsilon(x)} dx = \bar{J}_\varepsilon \cdot \int \frac{f^2(x)\varphi^2(x)}{f(x)\varphi(x) + \varepsilon(x)} dx = \\ &= (J[f] + \bar{\varepsilon}) \cdot \int f(x)\varphi(x) \frac{f^2(x)\varphi^2(x)\mu(dx)}{f(x)\varphi(x) + \varepsilon(x)} = \\ &= (J[f] + \bar{\varepsilon}) \cdot \int f(x)\varphi(x) \left(\frac{1}{1 + \varepsilon(x)/f(x)\varphi(x)} \right) \mu(dx) = \\ &= (J[f] + \bar{\varepsilon}) \cdot \int f(x)\varphi(x) \left[1 - \frac{\varepsilon(x)}{f(x)\varphi(x)} + \frac{\varepsilon^2(x)}{f^2(x)\varphi^2(x)} - O(\delta^3) \right] \leq \\ &\leq (J[f] + \bar{\varepsilon})(J[f] - \bar{\varepsilon}) + (J[f] + \bar{\varepsilon})(\delta^2 J[f] + O(\delta^3)), \end{aligned} \quad (3.1.14)$$

где обозначено

$$\delta = \sup_{(\text{mod } \mu)} \left| \frac{\varepsilon(x)}{f(x)\varphi(x)} \right|, \quad \bar{\varepsilon} = \int \varepsilon(x)\mu(dx).$$

Обозначение $\sup_{(\text{mod } \mu)}$ — указывает, что супремум вычисляется на носителе меры μ , который следует предполагать ограниченным. Заметим, что последнее неравенство использует положительность функции $f(x)\varphi(x)$.

Таким образом, отклонение на $\varepsilon(x)$ в оптимальной плотности вызывает, грубо говоря, отклонения порядка $\|\varepsilon\|^2$ у дисперсии оценки.

Включение особенности в плотность

Важной особенностью приемов понижения дисперсии является то, что их формальное применение в задачах, где оценка интеграла может не иметь второго момента, позволяет выбрать такой вариант оценки, где дисперсия конечна. Из теоремы 3.1 следует, что теоретически это можно сделать всегда, когда f абсолютно интегрируема относительно меры μ . Практическую сторону обычно поясняют на простейшем примере интеграла $\int_0^1 \frac{f(x)}{\sqrt{x}} dx$. Предполагается, что $f(x)$ отделена от нуля и ограничена на $[0, 1]$. Если использовать оценку

$$\varkappa_1 = \frac{1}{N} \sum_{j=1}^N \frac{f(\alpha_j)}{\sqrt{\alpha_j}},$$

где α_j равномерно распределены на $[0, 1]$, то легко видеть, что интеграл $\int_0^1 \frac{f^2(x)}{x} dx$, через который выражается ее второй момент, расходится. Однако достаточно использовать плотность $\frac{1}{2\sqrt{x}}$, чтобы получить оценку

$$\varkappa_2 = \frac{2}{\sqrt{N}} \sum_{j=1}^N f(\beta_j),$$

которая несмещена при независимых β_j , распределенных с этой плотностью, и имеет конечную дисперсию:

$$D\varkappa_2 = \frac{1}{N} \left(2 \int_0^1 \frac{f^2(x)}{\sqrt{x}} dx - \left(\int_0^1 \frac{f(x)}{\sqrt{x}} dx \right)^2 \right).$$

Прием такого рода называется включением особенности в плотность и широко используется на практике.

3.1.3. Интегрирование по части переменных

Предположим, что интеграл $J[f]$ представим в виде повторного

$$\int_{(\mathbf{x})} f(\mathbf{x}) \mu(d\mathbf{x}) = \int_{(\mathcal{Y})} \mu_1(dy) \int_{(\mathcal{Z})} \mu_2(dz) f(y, z), \quad (3.1.15)$$

то есть $\mathbf{x} = (y, z)$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, $\mathbf{x} = \mathcal{Y} \times \mathcal{Z}$ и $\mu = \mu_1 \otimes \mu_2$, и выполнено аналитическое интегрирование по переменной z , так что известна функция

$$F(y) = \int_{(\mathcal{Z})} \mu_2(dz) f(y, z). \quad (3.1.16)$$

Сравним дисперсии двух несмещенных оценок интеграла $J[f]$:

$$\varkappa_1 = \frac{1}{N} \sum_{j=1}^N f(Y_j, Z_j),$$

где случайные величины (векторы) Y_j распределены в соответствии с мерой μ_1 , а Z_j — в соответствии с мерой μ_2 , и

$$\varkappa'_1 = \frac{1}{N} \sum_{j=1}^N F(Y_j).$$

Поскольку их математические ожидания равны, достаточно сравнить вторые моменты. Рассмотрим разность

$$\begin{aligned} E\varkappa_1^2 - E(\varkappa'_1)^2 &= \frac{1}{N} \left[\int \mu_1(dy) \int \mu_2(dz) f^2(y, z) - \int \mu_1(dy) F^2(y) \right] = \\ &= \frac{1}{N} \int \mu_1(dy) \left[\int f^2(y, z) \mu_2(dz) - \left(\int f(y, z) \mu_2(dz) \right)^2 \right]. \end{aligned}$$

Но при каждом фиксированном y выражение в квадратных скобках неотрицательно (неравенство Коши–Буняковского). Отсюда получаем

$$E\varkappa_1^2 - E(\varkappa'_1)^2 \geq 0, \text{ то есть } D(\varkappa'_1) \leq D\varkappa_1. \quad (3.1.17)$$

Интегрирование по части переменных может только уменьшить дисперсию оценки.

Замечание. На этом примере особенно отчетливо видно, что уменьшение дисперсии не может служить самоцелью. Функция F , как правило, значительно сложнее функции f . Работа по вычислению значений F — трудоемкость вычисления оценки $\varkappa'_1$ — может значительно превосходить трудоемкость вычисления оценки $\varkappa_1$. В конкретных ситуациях вопрос о выгодности оценки $\varkappa'_1$ подлежит исследованию. Возможен, правда, случай, когда $D\varkappa_1 = +\infty$, а $D\varkappa'_1 < +\infty$; при этом, как правило, преимущество отдается оценке $\varkappa'_1$. $\square$

3.1.4. Метод расслоенной выборки

Метод основан на том, что интеграл $\int_{\mathfrak{X}} f(x)\mu(dx)$ может быть представлен в виде суммы интегралов

$$J[f] = \sum_{k=1}^m \int_{\mathfrak{X}_k} f(x)\mu(dx), \text{ где } \mathfrak{X} = \bigcup_{k=1}^m \mathfrak{X}_k \text{ и } \mathfrak{X}_k \cap \mathfrak{X}_l = \emptyset \text{ при } k \neq l.$$

Каждый из интегралов вычисляется с помощью оценки вида (3.1.8) с различным, вообще говоря, числом слагаемых при каждом k , то есть

$$J[f] \sim \sum_{k=1}^m \frac{\mu(\mathfrak{X}_k)}{n_k} \sum_{i_k=1}^{n_k} f(X_{i_k}^{(k)}), \quad (3.1.18)$$

где реализации $X_{i_k}^{(k)}$ имеют распределение

$$\mu_k(dx) = \begin{cases} \mu(dx)/\mu(\mathfrak{X}_k), & x \in \mathfrak{X}_k, \\ 0, & x \notin \mathfrak{X}_k. \end{cases}$$

Легко проверить, что второй момент этой оценки зависит от характера разбиения множества $\mathfrak{X}$ на подмножества $\mathfrak{X}_k$ и от чисел n_k . Возможна постановка задачи об оптимальном разбиении множества $\mathfrak{X}$ и выборе n_k ($\sum_{k=1}^m n_k = N$). Решение для $m = 2$ можно найти, например, в [35]. При фиксированном разбиении $\mathfrak{X} = \bigcup_{k=1}^m \mathfrak{X}_k$ наилучший результат получаем, когда n_k/N близки к $\int_{\mathfrak{X}_k} f(x)\mu(dx)$. Здесь внимательный читатель, возможно, усмотрит аналогию с методом существенной выборки.

Далее проанализируем частный случай, из которого следуют любопытные выводы. Пусть $m = 2$ и множество $\mathfrak{X}$ разделено произвольным образом на два непересекающихся подмножества равной меры: $\mu(\mathfrak{X}_1) = \mu(\mathfrak{X}_2) = 1/2$. Для вычисления интеграла используется оценка

$$J[f] \sim \frac{1}{2N} \sum_{j=1}^N \left(f(X_j^{(1)}) + f(X_j^{(2)}) \right) = \varkappa_3[f], \quad (3.1.19)$$

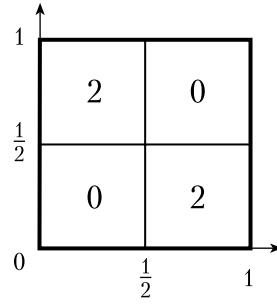


Рис. 3.2

где $X_j^{(1)}$ и $X_j^{(2)}$ — независимые при различных j реализации случайных величин, распределенных по законам $\mu_1(dx)$ и $\mu_2(dx)$ соответственно:

$$\mu_1(dx) = \begin{cases} 2\mu(dx), & x \in \mathfrak{X}_1, \\ 0, & x \notin \mathfrak{X}_1, \end{cases} \quad \mu_2(dx) = \begin{cases} 2\mu(dx), & x \in \mathfrak{X}_2, \\ 0, & x \notin \mathfrak{X}_2. \end{cases}$$

При одном и том же j , вообще говоря, $X_j^{(1)}$ и $X_j^{(2)}$ зависимы. Их совместная плотность по отношению к $\mu^2(dx, dy)$ есть

$$\varphi(x, y) = \begin{cases} 2, & x \in \mathfrak{X}_1, y \in \mathfrak{X}_2 \text{ или } x \in \mathfrak{X}_2, y \in \mathfrak{X}_1, \\ 0, & x \in \mathfrak{X}_1, y \in \mathfrak{X}_1 \text{ или } x \in \mathfrak{X}_2, y \in \mathfrak{X}_2. \end{cases}$$

В простейшем случае, когда $\mathfrak{X} = [0, 1]$, μ — мера Лебега, $\mathfrak{X}_1 = [0, 1/2]$, $\mathfrak{X}_2 = [1/2, 1]$, $\varphi(x, y)$ выглядит так, как это указано на рис. 3.2.

Вычислим

$$\begin{aligned} D\mathfrak{K}_3[f] &= \frac{1}{N} \left\{ \int_{\mathfrak{X}_1} \mu(dX) \int_{\mathfrak{X}_2} \mu(dY) 2 \left[\frac{1}{2} (f(X) + f(Y)) \right]^2 + \right. \\ &\quad \left. + \int_{\mathfrak{X}_2} \mu(dX) \int_{\mathfrak{X}_1} \mu(dY) 2 \left[\frac{1}{2} (f(X) + f(Y)) \right]^2 - (J[f])^2 \right\} = \\ &= \frac{1}{N} \left\{ \frac{1}{2} \int_{\mathfrak{X}} f^2(X) \mu(dx) + 2ab - (a+b)^2 \right\}, \end{aligned}$$

где обозначено $a = \int_{\mathfrak{X}_1} f(x) \mu(dx)$, $b = \int_{\mathfrak{X}_2} f(x) \mu(dx)$.

Поскольку $2ab - (a+b)^2 = ((a+b)^2 + (a-b)^2)/2$, то имеем

$$D\mathfrak{K}_3' = \frac{1}{N} \left\{ \frac{1}{2} \left(\int_{\mathfrak{X}} f^2(x) \mu(dx) - \left(\int_{\mathfrak{X}} f(x) \mu(dx) \right)^2 - (a-b)^2 \right) \right\}. \quad (3.1.20)$$

Теперь заметим, что при вычислении $J[f]$ с помощью оценки (3.1.8) с $2N$ испытаниями (то есть без расслоения) мы имели бы дисперсию

$$D\mathfrak{X}_1 = \frac{1}{2N} \left[\int f^2(x) \mu(dx) - \left(\int f(x) \mu(dx) \right)^2 \right].$$

Сравнивая это выражение с (3.1.20), получаем $D\mathfrak{X}_3 \leq D\mathfrak{X}_1$. Таким образом, расслоение (в этом частном случае, по крайней мере) никогда не увеличивает дисперсию, а уменьшает ее, если $a \neq b$ (при сохранении объема вычислительной работы).

Расслоение уменьшает случайность: можно гарантировать, что половина точек, в которых вычисляется f , будет находиться в $\mathfrak{X}_1$, а половина — в $\mathfrak{X}_2$. При отсутствии расслоения вероятность события $\{X_j^{(1)} \in \mathfrak{X}_1 \text{ и } X_j^{(2)} \in \mathfrak{X}_1\}$ положительна. Очевидно, разбивая $\mathfrak{X}$ на 4, 8 и т. д. частей, мы будем улучшать ситуацию и получим почти детерминированную процедуру. Сравнительно легко показать, что аналогичная ситуация имеет место при делении множества $\mathfrak{X}$ на любое число частей равной меры. И мы должны были бы изгнать случайность (метод Монте-Карло!), но если деление отрезка пополам не требует, в сущности, дополнительной вычислительной работы, то деление гиперкуба большой размерности на части равного объема — совсем непростая задача. Этой проблемы, очевидно, нет при использовании датчика случайных чисел. Существуют и другие преимущества подхода, основанного на случайном выборе. Об этом будет сказано в других разделах.

3.1.5. Случайные интерполяционно-квадратурные формулы

Пусть $\mathfrak{X}$ таково, что на нем могут быть определены линейно-независимые функции $\varphi_1(x), \dots, \varphi_m(x)$. Рассмотрим следующую задачу интерполирования функции $f(x)$. В $\mathfrak{X}$ выберем m различных точек $x_1, \dots, x_m$ (предполагается, что это можно сделать) и определим константы $C_1, \dots, C_m$ из условия

$$\sum_{j=1}^m C_j \varphi_j(x_l) = f(x_l) \pmod{\mu}, \quad l = 1, 2, \dots, m. \quad (3.1.21)$$

С помощью формул Крамера для интерполяционного многочлена

$$L[f, x] = \sum_{j=1}^m C_j \varphi_j(x),$$

удовлетворяющего (3.1.21), получаем выражение

$$L[f, x] = \frac{-1}{\det \|\varphi_i(x_j)\|_{i,j=1}^m} \begin{vmatrix} 0 & \varphi_1(x) & \dots & \varphi_m(x) \\ f(x_1) & \varphi_1(x_1) & \dots & \varphi_m(x_1) \\ \dots & \dots & \dots & \dots \\ f(x_m) & \varphi_1(x_m) & \dots & \varphi_m(x_m) \end{vmatrix} \quad (3.1.22)$$

в предположении, что $\det \|\varphi_i(x_j)\|_{i,j=1}^m$ отличен от нуля. Заметим, что проверка нужных нам свойств $L[f, x]$ может быть выполнена непосредственно (без формул Крамера). Действительно, разлагая определитель в числителе (3.1.22) по элементам первой строки, убеждаемся, что $L[f, x]$ имеет вид $\sum_{j=1}^m C_j \varphi_j(x)$. Подставляя вместо x значение x_k , имеем в качестве первой строки упомянутого определителя

$$0 \quad \varphi_1(x_k) \quad \dots \quad \varphi_m(x_k).$$

Вычитаем эту строку из $(k+1)$ -й строки и получаем в качестве $(k+1)$ -й строки

$$f(x_k) \quad 0 \quad \dots \quad 0.$$

Далее нужно раскрыть определитель по элементам этой строки, в оставшемся определителе поставить 1-ю строку на k -е место. Получаем $L[f, x_k] = f(x_k)$, $k = 1, 2, \dots, m$.

Предположим теперь, что $f(x)$ и $\varphi_l(x)$, $l = 1, 2, \dots, m$, интегрируемы с квадратом относительно меры μ . Тогда, не умаляя общности, можно считать, что φ_l ортонормированы. (Можно проверить даже, что процесс ортонормирования вызовет лишь деление C_j на константу нормировки.)

Если f и φ_l обладают некоторыми свойствами гладкости, то принято считать, что при удачном выборе точек x_k многочлен $L[f, x]$ будет приближать функцию f . Эти же соображения позволяют предполагать, что интеграл $\int \varphi_1(x) L[f, x] \mu(dx)$ будет приближать интеграл $\int \varphi_1(x) f(x) \mu(dx)$. Заметим, что при $\varphi_1(x) \equiv 1$ мы имеем интеграл $J[f]$. Кроме того, роль φ_1 может выполнять любая из φ_l , $l = 2, \dots, m$.

Используя равенство (3.1.22) и ортонормированность φ_l , имеем

$$\begin{aligned} & \int \varphi_1(x) L[f, x] \mu(dx) = \\ &= \frac{1}{\det \|\varphi_i(x_j)\|_{i,j=1}^m} \det \|f(x_k), \varphi_2(x_k), \dots, \varphi_m(x_k)\|_{k=1}^m. \end{aligned} \quad (3.1.23)$$

Разложение определителя в числителе (3.1.23) по элементам первого столбца дает

$$\frac{1}{\det \|\varphi_i(x_j)\|_{i,j=1}^m} \det \|f(x_k), \varphi_2(x_k), \dots, \varphi_m(x_k)\|_{k=1}^m =$$

$$= \sum_{k=1}^m A_k(x_1, \dots, x_m) f(x_k),$$

то есть интегралу соотносится кубатурная сумма

$$\int f(x) \varphi_1(x) \mu(dx) \sim \sum_{k=1}^m A_k(x_1, \dots, x_m) f(x_k). \quad (3.1.24)$$

Формулу (3.1.23) принято называть интерполяционно-кубатурной формулой.

Если, как мы говорили, можно ожидать, что сумма

$$K_m[f, x_1, \dots, x_m] = \sum_{k=1}^m A_k(x_1, \dots, x_m) f(x_k)$$

приближает искомый интеграл, то, выбирая точки x_l , $l = 1, \dots, m$, случайно и строя по этим точкам случайную кубатурную сумму, мы сможем получить инструмент для понижения дисперсии. Перейдем к строгой постановке задачи.

Пусть набор точек $Q = (x_1, \dots, x_m)$ выбирается случайно с плотностью $g(Q)$ относительно меры $\mu^m(dQ)$. Если определитель $\Delta(Q) = \det \|\varphi_i(x_j)\|_{i,j=1}^m$ отличен от нуля, то строится случайная кубатурная сумма $K[f, Q]$. Наша задача состоит в том, чтобы разумным образом выбрать $g(Q)$. Потребуем, чтобы выполнялось равенство

$$\mathbb{E} K[f, Q] = \int \mu^m(dQ) K[f, Q] g(Q) = \int \varphi_1(x) f(x) \mu(dx). \quad (3.1.25)$$

Пусть $m = 2$, μ — мера Лебега на отрезке $[0, 1]$, $\varphi_1 \equiv 1$, $\varphi_2(x) = \sqrt{12}(x - 1/2)$. Этот простой пример показывает, что нельзя выбирать x_1 и x_2 независимыми. Действительно, в этом случае

$$K[f, Q] = \frac{f(x_1)(x_2 - 1/2) - f(x_2)(x_1 - 1/2)}{x_2 - x_1}$$

и

$$\mathbb{E} K[f, Q] = \int_0^1 dx_1 \int_0^1 dx_2 \frac{f(x_1)(x_2 - 1/2) - f(x_2)(x_1 - 1/2)}{x_2 - x_1} g(x_1, x_2).$$

Интеграл имеет особенность вдоль диагонали квадрата при $x_2 = x_1$. Если f — произвольная μ -интегрируемая с квадратом функция, то по крайней мере для некоторых f интеграл будет расходиться, когда $g(x_1, x_2) = g_1(x_1)g_2(x_2)$, и конечно, выполнение равенства (3.1.25) в этом случае возможно лишь для некоторых функций.

Условием существования $K[f, Q]$ является отличие от нуля определителя $\Delta(Q)$. Следует ожидать, что $g(Q)$ будет содержать $\Delta(Q)$ в качестве множителя. Далее покажем, что

$$g(Q) = \text{Const} \cdot \Delta^2(Q)$$

обеспечивает выполнение равенства (3.1.25).

Следует отметить, что это не единственная плотность, удовлетворяющая условию (3.1.25). Подробный анализ проблемы выбора $g(Q)$ можно найти в книге [35]. Мы ограничимся здесь лишь простейшим частным случаем.

В дальнейшем изложении будем использовать лемму 3.1, являющуюся некоторым обобщением известных свойств интегралов от определителей.

Лемма 3.1. Пусть заданы два набора принадлежащих $\mathbf{L}^2(\mu)$ функций: $\varphi_1(x), \dots, \varphi_m(x)$ и $\psi_1(x), \dots, \psi_m(x)$. Справедливо равенство

$$\int \mu^m(dQ) \det \|\varphi_i(x_j)\|_{i,j=1}^m \times \\ \times \det \|\psi_k(x_l)\|_{k,l=1}^m = m! \det \|(\varphi_i, \psi_k)\|_{i,k=1}^m,$$

где

$$(\varphi_i, \psi_k) = \int \varphi_i(x) \psi_k(x) \mu(dx).$$

Доказательство. Разложим определитель $\det \|\varphi_i(x_j)\|_{i,j=1}^m$ (j — номер строки) по элементам первого столбца; соответствующий минор обозначим Δ_j :

$$\Delta_j = \begin{vmatrix} \varphi_2(x_1) & \varphi_3(x_1) & \dots & \varphi_m(x_1) \\ \dots & \dots & \dots & \dots \\ \varphi_2(x_{j-1}) & \varphi_3(x_{j-1}) & \dots & \varphi_m(x_{j-1}) \\ \varphi_2(x_{j+1}) & \varphi_3(x_{j+1}) & \dots & \varphi_m(x_{j+1}) \\ \dots & \dots & \dots & \dots \\ \varphi_2(x_m) & \varphi_3(x_m) & \dots & \varphi_m(x_m) \end{vmatrix}$$

Получим

$$\begin{aligned} I &= \int \mu^m(dQ) \sum_{j=1}^m (-1)^{j+1} \Delta_j \varphi_1(x_j) \det \|\psi_k(x_l)\|_{k,l=1}^m = \\ &= \sum_{j=1}^m \int \mu(dx_1) \dots \int \mu(dx_{j-1}) \int \mu(dx_{j+1}) \dots \int \mu(dx_m) \times \\ &\quad \times (-1)^{j+1} \Delta_j \int \mu(dx_j) \varphi(x_j) \det \|\psi_k(x_l)\|_{k,l=1}^m. \end{aligned}$$

Домножим строку определителя $\det \|\psi_k(x_l)\|_{k,l=1}^m$ с номером j на $\varphi_1(x_j)$ и проинтегрируем ее элементы. Перестановка проинтегрированной строки на первое место вызовет появление множителя $(-1)^{j-1}$. Если теперь изменить нумерацию первых $j-1$ переменных интегрирования во всех слагаемых, начиная со второго, обозначая x_2 вместо x_1 , x_3 — вместо x_2 и т. д., то получим

$$\begin{aligned} I &= m \int \mu(dx_2) \dots \int \mu(dx_m) \det \|\varphi_i(x_j)\|_{i,j=1}^m \times \\ &\quad \times \begin{vmatrix} (\varphi_1, \psi_1) & (\varphi_1, \psi_2) & \dots & (\varphi_1, \psi_m) \\ \psi_1(x_2) & \psi_2(x_2) & \dots & \psi_m(x_2) \\ \dots & \dots & \dots & \dots \\ \psi_1(x_m) & \psi_2(x_m) & \dots & \psi_m(x_m) \end{vmatrix} \end{aligned}$$

Таким образом, мы понизили на единицу порядок одного из определителей и проинтегрировали строку второго. Буквальное повторение этой операции, очевидно, приведет нас к желаемому результату. Лемма доказана. $\square$

С помощью этой леммы легко доказать следующую теорему.

Теорема 3.2. Пусть $\varphi_l(x)$, $l = 1, \dots, m$, — ортонормированная в $L^2(\mu)$ система функций. Тогда

- (1) $\frac{1}{m!} \Delta^2(Q)$ является плотностью по отношению к $\mu^m(dQ)$.
- (2) Если Q выбирается случайно в соответствии с этой плотностью, то справедливо равенство

$$\mathbb{E}K[f, Q] = \int \varphi_1(x) f(x) \mu(dx), \quad (3.1.26)$$

где $K[f, Q] = \frac{\Delta(f, Q)}{\Delta(Q)}$, $\Delta(f, Q) = \det \|f(x_k, \varphi_2(x_k), \dots, \varphi_m(x_k))\|_{k=1}^m$.

Доказательство. Утверждение (1) очевидно (матрица $\|(\varphi_i, \varphi_k)\|_{i,k=1}^m$ является единичной). Далее,

$$EK[f, Q] = \int \mu^m(dQ) \frac{\Delta(f, Q)}{\Delta(Q)} \cdot \frac{1}{m!} \Delta^2(Q) = \int \mu^m(dQ) \Delta(f, Q) \Delta(Q).$$

Для вычисления последнего интеграла используется утверждение леммы 3.1, если положить $\psi_1 = f$, $\psi_l = \varphi_l$, $l = 2, \dots, m$, и считать φ_l ортонормированными. Имеем

$$\begin{aligned} EK[f, Q] &= \begin{vmatrix} (f_1, \varphi_1) & (f_1, \varphi_2) & \dots & (f_1, \varphi_m) \\ 0 & 1 & \dots & 0 \\ \dots & \dots & 1 & \dots \\ 0 & \dots & \dots & 1 \end{vmatrix} = \\ &= (f_1, \varphi_1) = \int \varphi_1(x) f(x) \mu(dx), \end{aligned}$$

что и завершает доказательство. $\square$

Способ вычисления дисперсии $K[f, Q]$ зависит от свойств системы функций $\varphi_1, \dots, \varphi_m$.

Определение 3.1. Назовем систему функций $\{\varphi_l\}_1^m$ регулярной по отношению к мере μ , если она удовлетворяет условию

$$\mu^m\{Q : \Delta(Q) = 0\} = 0.$$

Примеры. (1) Функции $1, \sqrt{12}(x-1/2)$ образуют регулярную систему ($m = 2$). Определитель $\Delta(Q) = \sqrt{12}(x_2 - x_1)$ обращается в нуль на диагонали единичного квадрата, которая имеет μ^2 меру, равную нулю; μ — мера Лебега. Легко показать, что многочлены от одной и многих переменных регулярны в ограниченной области по отношению к мере Лебега.

(2) Примером нерегулярной системы по отношению к лебеговой мере на $(0, 1)$ является система ($m = 2$) функций

$$\varphi_1(x) \equiv 1, \quad \varphi_2(x) = \begin{cases} 1, & x \in [0, 1/2), \\ -1, & x \in [1/2, 1]. \end{cases} \quad (3.1.27)$$

Здесь $\Delta(Q) = (\varphi_2(x_2) - \varphi_2(x_1))$. Легко подсчитать, что $\mu^2\{x_1 x_2 : (\varphi_2(x_2) - \varphi_2(x_1)) = 0\} = 1/2$.

Функции (3.1.27) представляют собой две первые функции системы Хаара. Функции Хаара, сплайны и вейвлеты могут служить примерами нерегулярных систем. $\square$

Теперь можно доказать следующую теорему.

Теорема 3.3. В условиях теоремы 3.2 выполняется неравенство

$$\mathrm{DK}[f, Q] \leq \int \left[f(x) - \sum_{l=1}^m (f, \varphi_l) \varphi_l(x) \right]^2 \mu(dx). \quad (3.1.28)$$

При этом если система функций регулярна, то в (3.1.28) имеет место знак равенства.

Доказательство. Вычислим

$$\begin{aligned} \mathrm{EK}^2[f, Q] &= \int_{\{Q: \Delta(Q) \neq 0\}} \mu^m(dQ) \frac{\Delta^2(f, Q)}{\Delta^2(Q)} \cdot \frac{1}{m!} \Delta^2(Q) = \\ &= \frac{1}{m!} \int_{\{Q: \Delta(Q) \neq 0\}} \mu^m(dQ) \Delta^2(f, Q). \end{aligned}$$

Полученная величина меньше, а в случае регулярной системы функций — равна

$$\begin{aligned} \delta_m &= \frac{1}{m!} \int \mu^m(dQ) \Delta^2(f, Q) = \\ &= \frac{1}{m!} \int \mu^m(dQ) (\det \|f(x_k), \varphi_2(x_k), \dots, \varphi_m(x_k)\|)^2 = \\ &= \begin{vmatrix} (f, f) & (f, \varphi_2) & \dots & (f, \varphi_m) \\ (f, \varphi_2) & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ (f, \varphi_m) & 0 & \dots & 1 \end{vmatrix}. \end{aligned}$$

Последнее равенство имеет место на основании леммы 3.1, если в ее формулировке положить $\{\varphi_l\} = (f, \varphi_2, \dots, \varphi_m)$, $\{\psi_k\} = (f, \varphi_2, \dots, \varphi_l)$. Далее, разлагая определитель по элементам последней строки (или столбца), получим $\delta_m = \delta_{m-1} - (f, \varphi_m)^2$. Если учесть, что $\delta_1 = (f, f)$, то имеем $\delta_m = (f, f) - \sum_{k=2}^m (f, \varphi_k)^2$. Поскольку $\mathrm{DK}[f, Q] = \mathrm{EK}^2[f, Q] - (f, \varphi_1)^2$, то для случая регулярной системы функций:

$$\mathrm{DK}[f, Q] = (f, f) - \sum_{k=1}^m (f, \varphi_k)^2 = \int \left(f(x) - \sum_{k=1}^m (f, \varphi_k) \varphi_k(x) \right)^2 \mu(dx), \quad (3.1.29)$$

а для нерегулярных функций:

$$\mathrm{DK}[f, Q] \leq \int \left(f(x) - \sum_{k=1}^m (f, \varphi_k) \varphi_k(x) \right)^2 \mu(dx),$$

что и доказывает теорему. $\square$

Заметим, что в нерегулярном случае мы уже рассматривали пример вычисления дисперсии. Действительно, разбивая множество $\mathfrak{X}$ на два подмножества равной меры и полагая

$$\varphi_1(x) \equiv 1, \quad \varphi_2(x) = \begin{cases} 1, & x \in \mathfrak{X}_1, \\ -1, & x \in \mathfrak{X}_2, \end{cases}$$

мы, как легко проверить, получаем рассмотренный ранее случай расслоенной выборки.

Здесь

$$\frac{1}{2}\Delta^2(Q) = \begin{cases} 2, & \text{если } x_1 \in \mathfrak{X}_1, x_2 \in \mathfrak{X}_2 \text{ или } x_1 \in \mathfrak{X}_2, x_2 \in \mathfrak{X}_1, \\ 0, & \text{если } x_1 \text{ и } x_2 \text{ принадлежат одному из подмножеств,} \end{cases}$$

$$K[f, Q] = \frac{1}{2}(f(x_1) + f(x_2)),$$

а дисперсия оказывается равной

$$DK[f, Q] = \frac{1}{2}[(f, f) - (f, \varphi_1)^2 - (f, \varphi_2)^2].$$

Легко видеть, что выражение $(a - b)$, фигурирующее в (3.1.20), не что иное, как (f, φ_2) .

Необходимо сделать следующие замечания относительно интерполяционно-кубатурных формул и их использования.

Как показано в [35], распределение Δ^2 не является единственным, обеспечивающим несмещенность интерполяционно-кубатурной суммы. Более того, в нерегулярном случае всегда можно указать распределение, обеспечивающее меньшую дисперсию, чем это определено равенством (3.1.29).

Как расслоенная выборка, так и интерполяционно-кубатурные формулы используют зависимые в данной группе случайные величины. Более жесткая зависимость может улучшить результат (уменьшить дисперсию). В этой связи полезно изучать суммы $K[f, Q]$, у которых один из узлов (например, x_1) выбирается случайно, а другие являются его детерминированными функциями. Отдельные результаты вычисления таких формул можно найти в [35]. Самым известным примером является следующий [15]:

$$\int_{-1}^{+1} dx_1 \dots \int_{-1}^{+1} dx_s f(x) \sim f(Y) + f(-Y), \quad (3.1.30)$$

где Y равномерно распределен в гиперкубе $\{x_1, \dots, x_s : -1 \leq x_i \leq +1, i = 1, \dots, s\}$. Это так называемый “antithetic variate method” (метод противоположной переменной). Относительно дисперсии кубатурной суммы (3.1.30) см. также [35].

Моделирование распределения Δ^2 , если использовать метод отбора при большом m ($m \geq 10$), является достаточно трудоемкой процедурой. Вместе с тем использование кубатурных формул со случайными узлами, распределенными с плотностью Δ^2 , привлекательно своей универсальностью. Считалось, что трудность моделирования Δ^2 исключает практическое использование таких формул. Недавние исследования показали, однако, что применение широкого спектра приемов повышения результативности моделирования распределений могут привести к эффективным процедурам при достаточно больших m . Соответствующие алгоритмы с оценкой их трудоемкости описаны в работе [42].

3.2. Метод квазиМонте-Карло

Как будет ясно из дальнейшего изложения, так называемый метод квазиМонте-Карло является вполне детерминированным и больше связан с классической теорией кубатурных формул, чем с теорией вероятностей. Нельзя не отметить, однако, что вычислительная схема упомянутого метода имеет формальное сходство со схемами предыдущего раздела и в ряде случаев обеспечивает более быстрое убывание погрешности. Занимаясь методом Монте-Карло, следует иметь четкое представление о формально близком методе.

Нам понадобятся некоторые сведения из классической теории кубатурных формул.

3.2.1. Некоторые сведения из теории кубатурных формул

Пусть D_s — область s -мерного евклидова пространства ($D_s \subset \mathbf{R}^s$) и $f(X)$, $X \in D_s$ — функция, определенная и интегрируемая в этой области. Если $X_1, \dots, X_N$ — точки в D_s и $A_1, \dots, A_N$ — заданные константы, то формулу

$$\int_{D_s} f(X) dX \approx \sum_{j=1}^N A_j f(X_j) \quad (3.2.1)$$

называют кубатурной (квадратурной при $s = 1$) формулой, а сумму $S_N[f] = \sum_{j=1}^N A_j f(X_j)$ — кубатурной (квадратурной) суммой.

Величину

$$R_N[f] = \int_{D_s} f(X) dX - S_N[f] \quad (3.2.2)$$

называют остатком кубатурной формулы.

Если для некоторой функции φ , интегрируемой в D_s , имеет место равенство

$$R_N[\varphi] = 0, \quad (3.2.3)$$

то говорят, что формула (3.2.1) точна для этой функции.

В предыдущем параграфе мы рассматривали интерполяционно-кубатурные формулы, которые точны для m линейно независимых функций при $N = m$

$$R_m[\varphi_j] = 0, \quad j = 1, \dots, m. \quad (3.2.4)$$

При фиксированных $X_1, \dots, X_m$, удовлетворяющих условию

$$\det \|\varphi_j(X_j)\|_{i,j=1}^m \neq 0,$$

числа $A_1, \dots, A_m$ (коэффициенты формулы) определяются условиями (3.2.4) однозначно.

Очевидно, что за счет специального выбора узлов $X_1, \dots, X_m$ можно, вообще говоря, обеспечить выполнение равенств $R_m[\psi_k] = 0$ для некоторых функций ψ_k , $k = 1, 2, \dots, m_1$, отличных от φ_j . Формулы со специальным выбором узлов такого рода называют кубатурными формулами гауссова типа. При случайном выборе узлов (см. п. 3.1) речь идет о наложении на эти узлы детерминированных связей так, что некоторая группа l ($l < m$) узлов выбирается случайно, а остальные являются функциями выбранных узлов (например, см. (3.1.30)).

Проблема специального выбора узлов, в особенности когда φ_j и ψ_k есть полиномы, является одной из важных задач теории кубатурных формул.

Другой подход к построению кубатурных формул предполагает, что формула должна обслуживать некоторый класс $\mathcal{F}$ функций, то есть коэффициенты A_j и узлы X_j должны быть выбраны так, чтобы величина

$$\sup_{f \in \mathcal{F}} |R_m[f]|$$

была минимальной (существование супремума предполагается). Очевидно, что это более общая постановка задачи. Предыдущая может

рассматриваться в качестве ее частного случая ($\mathcal{F}$ есть множество линейных комбинаций φ_j и ψ_k).

Наиболее изучен случай, когда $\mathcal{F}$ является линейным нормированным пространством дифференцируемых функций. При этом

$$R_m[f] = R_m(f; A_1, \dots, A_m; X_1, \dots, X_m)$$

есть функционал на $\mathcal{F}$ и

$$\sup_{\|f\|_{\mathcal{F}}=1} |R_m[f]| = R_{\mathcal{F}}(A_1, \dots, A_m; X_1, \dots, X_m) \quad (3.2.5)$$

— его норма. Оптимальной формулой (если таковая существует) в классе функций $\mathcal{F}$ называется формула

$$\int_{D_s} f(X) dX \approx \sum A_j^0 f(X_j^0),$$

для которой

$$(A_1^0, \dots, A_m^0; X_1^0, \dots, X_m^0) = \underset{(A_j, X_j, j=1, \dots, m)}{\operatorname{argmin}} R_{\mathcal{F}}(A_1, \dots, A_m; X_1, \dots, X_m). \quad (3.2.6)$$

При изучении оптимальных формул имеются существенные трудности, иногда непреодолимые на современном уровне знаний. Первая трудность — это вычисление $\|R_N(A_i, X_i, i = 1, \dots, N)\|_{\mathcal{F}}$, а вторая — определение значений $A_i^0, X_i^0, i = 1, \dots, N$. Мы сталкиваемся, таким образом, с задачей отыскания минимума функций многих переменных, причем вычисление значений этой функции может быть весьма трудоемким. По этой причине значительный интерес представляет изучение различных свойств нормы функционала $R_N[f]$.

Наиболее исследованным представляется случай, когда $\mathcal{F}$ является множеством функций f , имеющих все частные производные заданного порядка m , то есть

$$f_{i_1, \dots, i_s}^m(x) = \frac{\partial^m f(x_1, \dots, x_s)}{\partial x_1^{i_1} \dots \partial x_s^{i_s}}, \quad (3.2.7)$$

где $i_1 + i_2 + \dots + i_s = m; 0 \leq i_l \leq m; l = 1, \dots, s$.

Бахваловым [28] получена оценка снизу нормы остатка. Точнее говоря, было показано, что если $f \in \mathbf{H}(m, a, \lambda)$, где $\mathbf{H}(m, a, \lambda)$ — класс функций, у которых частные производные (3.2.7) до порядка m включительно ограничены константой a , а производные порядка m удовлетворяют условию Гёльдера с показателем $\lambda, 0 < \lambda \leq 1$:

$$|f_{i_1, \dots, i_s}^m(X') - f_{i_1, \dots, i_s}^m(X'')| \leq a \sum_{q=1}^s |x'_q - x''_q|^\lambda,$$

то при любом способе вычисления интеграла с помощью квадратурной формулы (3.2.1) выполняется неравенство

$$\sup_{f \in \mathbf{H}(m, A, \lambda)} |R_N[f]| \geq Aa N^{-\frac{m+\lambda}{s}}.$$

Метод вычисления нормы функционала $R_N[f]$, когда $\mathcal{F}$ есть пространство $\mathbf{L}_2^{(m)}$ или $\mathbf{W}_m^2$, разработан С. Л. Соболевым [69]. Им же получен ряд глубоких результатов относительно свойств этой нормы. Тем не менее, эффективных способов решения проблемы (3.2.6) мы сегодня не знаем. Задача может быть решена при $s = 1$ для очень ограниченного перечня частных случаев.

Было также показано [29], что рандомизация (случайный выбор кубатурной формулы из некоторого множества формул) позволяет получить неравенство

$$\sup_{f \in \mathbf{H}(m, A, \lambda)} \mathbf{E} |R_N[f]| \geq Aa N^{-(m+\lambda)/s-1/2}.$$

Известны примеры, где остаток при рандомизации действительно уменьшается в $\sqrt{N}$ раз (расслоение).

Значительный интерес представляют собой результаты, касающиеся асимптотики нормы остатка (3.2.5) при $m \rightarrow \infty$, когда $\mathcal{F}$ является широким классом функций. Результаты такого рода получены в теории чисел в связи с исследованием отклонения (discrepancy) и являются основой метода квазиМонте-Карло.

3.2.2. Отклонение и его связь с нормой остатка

Пусть $s = 1$ и $D = [0, 1]$. Если $f(x)$ имеет первую производную, то

$$f(x) = f(0) + \int_0^x f'(t) dt. \quad (3.2.8)$$

Обозначим

$$\Theta(x) = 1 - \mathbf{E}(x) = \begin{cases} 0, & x > 0, \\ 1/2, & x = 0, \\ 1, & x < 0. \end{cases}$$

Тогда при $A_j = 1/N$, $j = 1, \dots, N$,

$$R[f] = \int_0^1 f'(t) \left(1 - t - \frac{1}{N} \sum_{i=1}^N \mathbf{E}(x_i - t)\right) dt,$$

что легко получить из (3.2.8). Нетрудно также увидеть, что $R[f] = \int_0^1 f'(t) [-t + \frac{1}{N} \sum_{i=1}^N \Theta(x_i - t)] dt$. В математической статистике функция $\frac{1}{N} \sum_{i=1}^N \Theta(x_i - t)$ именуется эмпирической функцией распределения случайной величины, имеющей реализации $x_1, x_2, \dots, x_N$. Кроме того, $\frac{1}{N} \sum_{i=1}^N \Theta(x_i - t) = \frac{1}{N} \sum_{i=1}^N \chi_{(-\infty, X_i)}(t) = \frac{1}{N} \times (\text{число точек из множества } \{x_1, \dots, x_N\}, \text{ координаты которых не превосходят } t)$. Величина же

$$\sup_{t \in (0,1)} N \left| t - \frac{1}{N} \sum_{i=1}^N \Theta(x_i - t) \right| \quad (3.2.9)$$

носит название отклонения (discrepancy) и принимается в теории чисел в качестве меры неравномерности распределения множества точек $x_1, \dots, x_N$. Предполагается $0 \leq x_i \leq 1$. Полезно помнить, что t на промежутке $(0,1)$ есть $P(\alpha \leq t)$ — функция распределения равномерно распределенной случайной величины α . В математической статистике величина (3.2.9), деленная на N , фигурирует под названием критерия Колмогорова.

Теперь можно заметить, что для функции f , имеющей суммируемую производную f' , справедлива точная оценка

$$|R_N[f]| \leq \int_0^1 |f'(t)| dt \cdot \sup_{t \in [0,1]} \left| t - \frac{1}{N} \sum \Theta(X_i - t) \right|. \quad (3.2.10)$$

В многомерном случае имеет место похожий результат, известный как неравенство Коксмы–Хлавки [24]. Однако его нельзя получить непосредственно из разложения функции f в ряд Тейлора в окрестности точки $(0,0)$, как мы могли это сделать в одномерном случае. Возникающие здесь вопросы полезно пояснить для двумерного случая. Общий случай довольно громоздок, и мы воспользуемся известным и уже упоминавшимся неравенством Коксмы–Хлавки.

Итак, пусть $s = 2$, в единичном квадрате имеется множество N точек $(x_1, y_1), \dots, (x_N, y_N)$ и задана функция $f(x, y)$, имеющая смешанную производную $\frac{\partial^2 f}{\partial x \partial y}$ и первые производные $\frac{\partial f}{\partial x}$ и $\frac{\partial f}{\partial y}$. Относительно последних считаем, что они заданы на сторонах квадрата (например, при $0 \leq x \leq 1$, $y = 0$ и $0 \leq y \leq 1$, $x = 0$ соответственно).

В этом случае мы, как и ранее, можем использовать представление

$$f(x, y) = f(0, 0) + \int_0^x f'_x(u, 0) du + \int_0^y f'_y(v, 0) dv + \int_0^x \int_0^y f''_{xy}(u, v) du dv,$$

которое приводит нас к представлению остатка:

$$\begin{aligned}
& \int_0^1 \int_0^1 f(x, y) dx dy - \frac{1}{N} \sum_{j=1}^N f(x_j, y_j) = \\
& = \int_0^1 f'_x(u, 0) \left[1 - u - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(x_j - u) \right] du + \\
& + \int_0^1 f'_y(0, v) \left[1 - v - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(y_j - v) \right] dv + \\
& + \int_0^1 \int_0^1 f''_{xy}(u, v) \left[(1 - u)(1 - v) - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(x_j - u) \mathbb{E}(y_j - v) \right] du dv.
\end{aligned} \tag{3.2.11}$$

Пусть теперь $\mathcal{F}$ есть класс функций f с нормой

$$\|f\|_{\mathcal{F}} = \int_0^1 |f'_x(u, 0)| du + \int_0^1 |f'_y(0, v)| dv + \int_0^1 \int_0^1 |f''_{xy}(u, v)| du dv. \tag{3.2.12}$$

В этом классе имеем точную оценку остатка $\|R_N[f]\|$:

$$\begin{aligned}
\|R_N[f]\| & \leq \|f\|_{\mathcal{F}} \cdot \max \left(\sup_u \left| 1 - u - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(x_j - u) \right|, \right. \\
& \quad \left. \sup_v \left| 1 - v - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(y_j - v) \right|, \right. \\
& \quad \left. \sup_{u, v} \left| (1 - u)(1 - v) - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(x_j - u) \mathbb{E}(y_j - v) \right| \right) = \\
& = \|f\|_{\mathcal{F}} \cdot \sup_{u, v} \left| (1 - u)(1 - v) - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(x_j - u) \mathbb{E}(y_j - v) \right|.
\end{aligned} \tag{3.2.13}$$

Последнее равенство следует из того, что первые два супремума под символом $\max$ могут быть получены из третьего, если в последнем положить $v = 0$ и $u = 0$ соответственно.

Неравенство является точным: существует f из $\mathcal{F}$, для которой в (3.2.13) достигается знак равенства. Такую функцию легко указать в явной форме.

Таким образом, величина

$$\sup_{u,v} \left| (1-u)(1-v) - \frac{1}{N} \sum_{j=1}^N \mathbb{E}(x_j - u) \mathbb{E}(y_j - v) \right| \quad (3.2.14)$$

вполне характеризует множество точек (x_j, y_j) , но, как легко проверить, она не совпадает с двумерным отклонением, которое должно иметь вид

$$\begin{aligned} & \sup_{u,v} N \left| uv - \frac{1}{N} \sum_{j=1}^N \Theta(x_j - u) \Theta(y_j - v) \right| = \\ & = \sup_{u,v} N \left| uv - \frac{1}{N} \sum_{j=1}^N (1 - \mathbb{E}(x_j - u))(1 - \mathbb{E}(y_j - v)) \right|, \end{aligned} \quad (3.2.15)$$

ср. (3.2.14).

Вернемся к (3.2.11). После элементарных преобразований имеем

$$\begin{aligned} R_N[f] &= \int_0^1 f'_x(u, 0) \left[-u + \frac{1}{N} \sum_{j=1}^N \Theta(x_j - u) \right] du + \int_0^1 f'_y(0, v) \left[-v + \right. \\ &+ \left. \frac{1}{N} \sum_{j=1}^N \Theta(y_j - v) \right] dv + \int_0^1 \int_0^1 f''_{xy}(u, v) \left[uv - \frac{1}{N} \sum_{j=1}^N \Theta(x_j - u) \Theta(y_j - v) - \right. \\ &\left. - u + \frac{1}{N} \sum_{j=1}^N \Theta(x_j - u) - v + \frac{1}{N} \sum_{j=1}^N \Theta(y_j - v) \right] du dv. \end{aligned} \quad (3.2.16)$$

Выполняя интегрирование второй производной, после приведения подобных членов получим выражение *)

$$\begin{aligned} R_N[f] &= \int_0^1 f'_x(u, 1) \left[-u + \frac{1}{N} \sum_{j=1}^N \Theta(x_j - u) \right] du + \\ &+ \int_0^1 f'_y(1, v) \left[-v + \frac{1}{N} \sum_{j=1}^N \Theta(y_j - v) \right] dv + \\ &+ \int_0^1 \int_0^1 f''_{xy}(u, v) \left[uv - \frac{1}{N} \sum_{j=1}^N \Theta(x_j - u) \Theta(y_j - v) \right] du dv. \end{aligned}$$

*) Легко проверить, что это представление можно получить из разложения в ряд Тейлора в окрестности точки $(1,1)$; см. [70].

Таким образом, в классе $\mathcal{F}'$ функций f , для которых ограничена сумма

$$\int_0^1 |f'_x(u, 1)| du + \int_0^1 |f'_y(1, v)| dv + \int_0^1 \int_0^1 \left| \frac{\partial^2 f}{\partial x \partial y}(u, v) \right| du dv = \|f\|_{\mathcal{F}'},$$

имеем

$$\begin{aligned} \|R_N[f]\| &\leq \|f\|_{\mathcal{F}'} \max \left(\sup_u \left| u - \frac{1}{N} \sum \Theta(x_j - u) \right|, \right. \\ &\sup_v \left| v - \frac{1}{N} \sum \Theta(y_j - v) \right|, \left. \sup_{u,v} \left| uv - \frac{1}{N} \sum \Theta(x_j - u) \Theta(y_j - v) \right| \right) = \\ &= \|f\|_{\mathcal{F}'} \cdot \sup_{u,v} \left| uv - \frac{1}{N} \sum \Theta(x_j - u) \Theta(y_j - v) \right|. \end{aligned} \quad (3.2.17)$$

Последнее равенство обосновывается аналогично (3.2.13). Таким образом, отклонение (3.2.15), деленное на N , является нормой функционала остатка кубатурной формулы с равными коэффициентами в классе функций, имеющих производные f'_x, f'_y, f''_{xy} (первые две должны существовать на соответствующих сторонах квадрата).

Техника, связанная с разложением в ряд Тейлора, позволяет достигать результатов и в многомерном случае. Похожим способом можно получить и упоминавшееся неравенство Коксмы–Хлавки, которое утверждает, что величина D_N/N , где

$$D_N = \sup_{x \in D_s} N \left| \prod_{i=1}^s x^{(i)} - \sum_{j=1}^N \prod_{i=1}^s \Theta(x_j^{(i)} - x^{(i)}) \right|, \quad X = (x^{(1)}, \dots, x^{(s)}), \quad (3.2.18)$$

является нормой остатка в классе функций ограниченной вариации в смысле Харди–Краузе. В двумерном случае это класс функций, для которого сумма

$$\int_0^1 d|f(x, 1)| + \int_0^1 d|f(1, y)| + \int_0^1 \int_0^1 d^2|f(x, y)|$$

ограничена константой M . Подробное изложение вопросов, связанных с этим неравенством, можно найти в [24], с. 165.

Укажем в качестве примера две широко известные в настоящее время конструкции последовательностей векторов $X_1, X_2, \dots, X_N, \dots$, для которых $R_N[f] \sim O\left(\frac{\ln^s N}{N}\right)$.

Примеры. (1) Последовательности Холтона. Пусть r, n и m — натуральные числа, и n имеет следующее представление в r -ичной системе счисления $n = a_m a_{m-1} \dots a_2 a_1$, $0 \leq a_i \leq r - 1$. Последовательность чисел $p_r(n), n = 1, 2, \dots$, определяется следующим образом:

$$p_r(n) = \sum_{i=1}^m a_i \cdot r^{-i}.$$

Последовательностью Холтона называется последовательность векторов в единичном гиперкубе D_s :

$$X(n) = (p_{r_1}(n), p_{r_2}(n), \dots, p_{r_s}(n)),$$

где $r_1, r_2, \dots, r_s$ — попарно взаимно простые числа.

(2) $\Lambda\Pi_r$ — последовательности Соболя. Это последовательность векторов в D_s :

$$Q_n = (q_{n,1}, \dots, q_{n,s}).$$

Алгоритм вычисления $q_{i,j}$ следующий. Если в двоичной системе счисления $n = l_m l_{m-1} \dots l_2 l_1$, $l_i = \{0, 1\}$, то для всех $j = 1, 2, \dots, s$

$$q_{n,j} = l_1 v_j^{(1)} * l_2 v_j^{(2)} * \dots * l_m v_j^{(m)}.$$

Звездочкой (*) обозначена операция поразрядного сложения по модулю 2 в двоичной системе счисления, а $v_j^{(i)}$ — заданные числа, которые в виде таблиц приведены в ряде руководств (см. приложение 3). $\square$

Таким образом, последовательности Холтона [13] и Соболя [70] определяют последовательность кубатурных формул с равными коэффициентами. Эти формулы обеспечивают в весьма широком классе функций (функций ограниченной вариации) убывание остатка порядка $O(\frac{\ln^s N}{N})$. Сравнительно простой алгоритм построения этих формул позволяет говорить об их использовании в схеме метода Монте-Карло: $p_r(n)$ и $q_{n,j}$ играют роль псевдослучайных чисел.

Имеются примеры удачного применения метода *квази*Монте-Карло. Вместе с тем нужно иметь в виду следующее:

- Как мы видим, точки Соболя настроены на определенную размерность вычисляемого интеграла. Координаты этих точек нельзя брать в произвольном порядке.
- Если рассматривать $p_2(n)$ и $q_{n,j}$ как реализации некоторых случайных величин, то оказывается, что эти величины, вообще говоря, коррелированы. Использование их при имитационном моделировании реальных явлений может приводить (особенно на уровне

вторых моментов) к грубым ошибкам. Конечно, числа, порождаемые мультипликаторным датчиком, также коррелированы, но с ростом мультипликатора корреляция убывает.

- В схеме *квази*Монте-Карло невозможно осуществить эмпирическую оценку погрешности с помощью дисперсии.

В целом, на стыке Монте-Карло и *квази*Монте-Карло требуются интенсивные дополнительные исследования.

3.3. Библиографическое дополнение к главе 3

В главе описаны традиционные приемы оценки интегралов методом Монте-Карло. Метод рассматривается, в первую очередь, как средство вычисления интегралов по вероятностной мере. Что касается кратных интегралов по единичному гиперкубу, то эта задача интересна постольку, поскольку к ней сводится, часто приближенно, предыдущая. По этой причине вопросы интегрирования гладких функций практически не обсуждаются.

С рандомизованными алгоритмами интегрирования гладких функций можно познакомиться в ставших классическими работах Бахвалова (например, [28]). Мы старались не использовать априорную информацию о гладкости подынтегральной функции. Незатронутыми оказались относящиеся к этому кругу идей последовательные процедуры интегрирования. В этой связи можно сослаться на работу Холтона [13]. Другие ссылки имеются в книге [24].

Отдельным важным разделом можно считать вычисление интегралов, зависящих от параметра. Предложенный Ченцовым и Фроловым метод зависимых испытаний подробно изложен в ряде популярных руководств (например, [43], ранее в [80]).

Упоминание о методе *квази*Монте-Карло необходимо, чтобы обсудить его отличие от метода Монте-Карло и в какой-то мере указать круг задач, где его целесообразно использовать. Следует отдавать себе отчет в том, что минимизация отклонения является по сути задачей классической теории кубатурных формул, в случае когда подынтегральная функция имеет возможно малую гладкость. О применении метода *квази*Монте-Карло в задачах, связанных с решением уравнений, см. также в гл. 5 этой книги.

Отметим, что иногда к методам Монте-Карло относят все теоретико-числовые и близкие к ним методы приближенного интегрирования. Как правило, исследования в области этих методов развивают и обобщают классические результаты Коробова [55]. При этом рассматриваются более узкие классы функций, чем класс функций ограниченной в смысле Харди–Крауде вариации.

Глава 4

Вычисление суммы ряда Неймана

4.1. Несмещенные оценки

Далее мы будем рассматривать уравнения 2-го рода:

$$\varphi(x) = \int k(x, y)\varphi(y)\mu(dy) + f(x) \quad (\text{mod } \mu), \quad (4.1.1)$$

где $k(x, y)$ и $f(x)$ — функции, заданные на носителях $\mathfrak{X}$ и $\mathfrak{X} \times \mathfrak{X}$ мер $\mu(dx)$ и $\mu(dx) \otimes \mu(dy)$ соответственно. Считается также, что они μ - и $\mu \otimes \mu$ -измеримы. Для простоты μ полагается вероятностной мерой. Все результаты этой главы могут быть перенесены на случай σ -конечных мер. Обозначение $(\text{mod } \mu)$ в (4.1.1) и далее означает, что равенство имеет место для $x \in \mathfrak{X}$.

В главе 2 обсуждался один из простейших вероятностных алгоритмов решения системы линейных алгебраических уравнений. Полезно помнить, что уравнение (4.1.1) есть система линейных алгебраических уравнений, если μ — дискретная мера, сосредоточенная в конечном множестве точек $x_1, \dots, x_n$ с весами $p_1, \dots, p_n$ соответственно.

В этом случае имеем

$$\varphi(x) = \sum_{j=1}^n p_j k(x, x_j) \varphi(x_j) + f(x) \quad (\text{mod } \mu). \quad (4.1.2)$$

Здесь обозначение $(\text{mod } \mu)$ подразумевает, что соотношение (4.1.2) выполняется в точках $x_1, \dots, x_n$, то есть может быть записано в виде

$$\varphi(x_i) = \sum_{j=1}^n p_j k(x_i, x_j) \varphi(x_j) + f(x_i), \quad i = 1, \dots, n, \quad (4.1.3)$$

или, если обозначить $\varphi(x_i) = z_i$, $p_j k(x_i, x_j) = a_{i,j}$, $f(x_i) = b_i$, то в виде (ср. (2.2.25))

$$z_i = \sum_{j=1}^n a_{i,j} z_j + b_i, \quad i = 1, \dots, n.$$

Обсуждаемые ниже результаты будут справедливы, таким образом, для уравнений Фредгольма, для систем линейных алгебраических уравнений и для более общих типов уравнений, порождаемых конкретными классами мер (см. начало главы 3).

Во всех случаях будем предполагать, что для некоторого класса $\mathcal{H}$ функций h сходится ряд

$$\sum_{l=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_l) |h(x_0)k(x_0, x_1) \dots k(x_{l-1}, x_l)f(x_l)| < \infty, \quad (4.1.4)$$

где положено $k(x_{-1}, x_0) \equiv 1$.

Сумма (4.1.4) есть скалярное произведение $(|h|, \bar{\varphi})$, где $\bar{\varphi}$ является итерационным решением уравнения

$$\bar{\varphi} = \bar{K}\bar{\varphi} + |f|, \quad \bar{K}\varphi[x] = \int |k(x, y)|\varphi(y)\mu(dy). \quad (4.1.5)$$

В этом случае также, очевидно, имеет место сходимость итерационного процесса

$$\begin{aligned} \varphi_{n+1}(x) &= \int k(x, y)\varphi_n(y)\mu(dy) + f(x), \quad n = 0, 1, 2, \dots \\ \varphi_0(x) &= f(x). \end{aligned} \quad (4.1.6)$$

Отметим, что можно уточнять, о сходимости в метрике какого функционального пространства идет речь; однако в нашем случае мы просто вычисляем сумму числового ряда. Достаточно предполагать, что сходится числовая последовательность функционалов (h, φ_n) для h из заданного множества $\mathcal{H}$. То есть предполагается, что имеет место представление

$$(h, \varphi) = \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_{\tau}) h(x_0)k(x_0, x_1) \dots k(x_{\tau-1}, x_{\tau})f(x_{\tau}), \quad (4.1.7)$$

где $h \in \mathcal{H}$, а φ является итерационным решением (4.1.1) в слабом (указанном выше) смысле.

Как и в рассмотренном нами в п. 2.2 случае систем линейных алгебраических уравнений, уравнению (4.1.1) сопоставляется цепь Маркова с множеством состояний $\mathfrak{X}$ и дискретным временем $t = 0, 1, 2, \dots$. Цепь задается плотностью начального распределения $p^0(x)$ по отношению к мере μ и плотностью перехода из состояния x в состояние

$y : p(x \rightarrow y)$ также по отношению к мере μ . Относительно плотности $p(x \rightarrow y)$ предполагается, что она субстохастична

$$\int p(x \rightarrow y)\mu(dy) = 1 - g(x), \quad 0 \leq g(x) < 1, \quad (4.1.8)$$

и цепь Маркова, ею определяемая, такова, что почти все ее траектории конечны.

Моделирование этой цепи в соответствии с п. 2.1 осуществляется с помощью следующего алгоритма:

1. Моделируется плотность $p^0(x)$. Вычисляется x_0 .
2. При $k > 0$ вычисляется $g(x_{k-1})$. Проверяется неравенство $\alpha < g(x_{k-1})$. При его вычислении полагаем $k = \tau$ и заканчиваем вычисления. Переход к 4.
3. Если $\alpha > g(x_{k-1})$, то моделируется плотность $p(x_{k-1} \rightarrow x)$. Вычисляется x_k , k заменяется на $k + 1$. Переход к 2.
4. Переход к моделированию следующей траектории (переход к 1), если необходимое число N траекторий еще не промоделировано.

Здесь α обозначает независимую реализацию равномерно распределенной случайной величины (очередное псевдослучайное число), τ — момент обрыва траектории. Результатом является множество N независимых траекторий $\omega_\tau = x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_\tau$ марковской цепи.

Очевидно, плотность распределения по отношению к $\mu^{\tau+1}$ такой траектории есть

$$p^0(x_0)p(x_0 \rightarrow x_1) \dots p(x_{\tau-1} \rightarrow x_\tau)g(x_\tau), \quad (4.1.9)$$

$\mu^{\tau+1} = \bigotimes_{j=0}^{\tau} \mu$. Совокупность этих плотностей определяет вероятностную меру $P(d\omega_\tau)$ на множестве траекторий. Поскольку мы предположили, что почти все траектории ω конечны, то должно выполняться равенство

$$\sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_\tau) p^0(x_0)p(x_0 \rightarrow x_1) \dots p(x_{\tau-1} \rightarrow x_\tau)g(x_\tau) = 1. \quad (4.1.10)$$

Данное равенство также легко можно получить, если оператор $\mathcal{P}$, определяемый равенством $\mathcal{P}\chi[x] = \int p(x \rightarrow y)\chi(y)\mu(dy) \pmod{\mu}$, является оператором сжатия.

Действительно, уравнение $\chi = \mathcal{P}\chi + g$ имеет в этом случае, вне всякого сомнения, своим единственным решением $\chi = 1 \pmod{\mu}$,

которое также имеет представление

$$\chi(x) = \sum_{\tau=0}^{\infty} \int \mu(dx_1) p(x_0 \rightarrow x_1) \dots p(x_{\tau-1} \rightarrow x_{\tau}) \pmod{\mu},$$

$$p(x_{-1}, x_0) = 1.$$

И в этом случае равенство $(\chi, p^0) = (1, p^0) = 1$ равносильно (4.1.10). Если же $\mathfrak{J}(\omega_{\tau})$ — функция траектории, интегрируемая по отношению к мере $P(d\omega_{\tau})$, то имеем

$$\begin{aligned} E\bar{\mathfrak{J}}(\omega_{\tau}) &= \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_{\tau}) \bar{\mathfrak{J}}(\omega_{\tau}) p^0(x_0) p(x_0 \rightarrow x_1) \dots \\ &\dots p(x_{\tau-1} \rightarrow x_{\tau}) g(x_{\tau}). \end{aligned} \quad (4.1.11)$$

Дальнейшее сравнение (4.1.7) с (4.1.10) позволяет доказать следующую теорему.

Теорема 4.1. Для того чтобы оценка

$$\mathfrak{J}_a(\omega_{\tau}) = \frac{h(x_0)k(x_0, x_1) \dots k(x_{\tau-1}, x_{\tau})f(x_{\tau})}{p^0(x_0)p(x_0 \rightarrow x_1) \dots p(x_{\tau-1} \rightarrow x_{\tau})g(x_{\tau})} \quad (4.1.12)$$

была несмещенной оценкой функционала (φ, h) , необходимо и достаточно сходимости ряда (4.1.4) и выполнения условий согласования:

- (1) $p^0(x) > 0$ для тех $x \pmod{\mu}$, для которых $h(x) \neq 0 \pmod{\mu}$,
- (2) $p(x \rightarrow y) > 0$ для тех $(x, y) \pmod{\mu^2}$, для которых $k(x, y) \neq 0 \pmod{\mu^2}$,
- (3) $g(x) > 0$ для тех $x \pmod{\mu}$, для которых $f(x) \neq 0 \pmod{\mu}$

(4.1.13)

при выполнении (4.1.10). $\square$

При доказательстве следует воспользоваться равенством (4.1.10) и учесть, что (4.1.4) является необходимым и достаточным условием интегрируемости оценки (4.1.12) на множестве траекторий.

Таким образом, стохастический (один из возможных) алгоритм решения уравнения (4.1.1) состоит в моделировании N независимых траекторий цепи Маркова, удовлетворяющей довольно общим условиям согласования, и вычисления на этих траекториях функции (4.1.12). Среднее арифметическое полученных результатов служит оценкой функционала (φ, h) .

Алгоритм обладает естественным параллелизмом (каждая траектория может вычисляться отдельным процессором) и простотой программной реализации.

Вопрос, возникающий в связи с изучением стохастических алгоритмов решения уравнений (4.1.1), состоит в следующем: будет ли (4.1.12) единственной функцией траекторий цепи Маркова, являющейся несмещенной оценкой функционала (φ, h) ?

Ответ сугубо отрицателен. Существует бесконечно много таких функций. (Подробно об этом см. в [35], гл. 6.)

Мы остановимся лишь на несложных в вычислительном отношении оценках. Несомненно, простота вычислений является немаловажным преимуществом. Оценка (4.1.12) достаточно проста. Почти столь же простой является оценка $\mathfrak{J}(\omega_\tau)$, имеющая вид

$$\mathfrak{J}_c(\omega_\tau) = \sum_{l=0}^{\tau} \frac{h_0 k_{0,1} \dots k_{l-1,l} f_l}{p_0^0 p_{0,1} \dots p_{l-1,l}}, \quad (4.1.14)$$

где обозначено $h_0 = h(x_0)$, $k_{i,j} = k(x_i, x_j)$, $f_l = f(x_l)$, $p_0^0 = p^0(x_0)$, $p_{i,j} = p(x_i \rightarrow x_j)$. В соответствии с традицией, возникшей в задачах переноса излучения через вещество, оценку (4.1.13) часто называют “оценкой по поглощению” (absorbtion), а оценку (4.1.14) — “оценкой по столкновениям” (collision). Важным преимуществом оценки (4.1.14) является то, что условия согласования для нее менее ограничительны.

Теорема 4.2. Оценка (4.1.14) является несмещенной оценкой функционала (h, φ) для всех h, K, f , удовлетворяющих условию (4.1.14), и p^0, p , удовлетворяющих (4.1.10) и следующим условиям согласования:

- (1) $p^0(x) > 0$ для тех $x \pmod{\mu}$, для которых $h(x) \neq 0 \pmod{\mu}$,
- (2) $p(x \rightarrow y) > 0$ для тех $(x, y) \pmod{\mu^2}$, для которых $k(x, y) \neq 0 \pmod{\mu^2}$

(4.1.15)

Доказательство. В соответствии с определением математического ожидания, имеем

$$\begin{aligned} E\mathfrak{J}_c(\omega_\tau) &= \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_\tau) \sum_{l=0}^{\tau} \frac{h_0 k_{0,1} \dots k_{l-1,l} f_l}{p_0^0 p_{0,1} \dots p_{l-1,l}} \times \\ &\quad \times p_{0,1}^0 \dots p_{\tau-1,\tau} g_\tau = \sum_{\tau=0}^{\infty} \sum_{l=0}^{\tau} \int \mu(dx_0) \dots \\ &\quad \dots \int \mu(dx_l) h_0 k_{0,1} \dots k_{l-1,l} f_l \int \mu(dx_{l+1}) \dots \int \mu(dx_\tau) p_{l,l+1} \dots p_{\tau-1,\tau} g_\tau, \end{aligned}$$

где, как и ранее, полагается $k_{-1,0} \equiv 1$.

При выполнении условий (4.1.4) (и (4.1.10)) мы можем изменять порядок суммирования, а также суммирования и интегрирования в двойной сумме.

Имеем

$$\begin{aligned} E\mathfrak{J}_c(\omega_\tau) = & \sum_{l=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_l) h_0 k_{0,1} \dots k_{l-1,l} f_l \sum_{\tau=l}^{\infty} \int \mu(dx_{l+1}) \dots \\ & \dots \int \mu(dx_\tau) p_{l,l+1} \dots p_{\tau-1,\tau} g_\tau. \end{aligned}$$

Но вторая сумма (по τ от l до ∞) изменением нумерации переменных может быть приведена к сумме (4.1.10), которая равна 1. Отсюда и следует утверждение теоремы. $\square$

Множество несмещенных оценок функционала (φ, h) можно тривиальным образом расширить.

Во-первых, при выполнении условий согласования (4.1.13) любая линейная комбинация $\alpha \mathfrak{J}_c(\omega_\tau) + \beta \mathfrak{J}_a(\omega_\tau)$ при $\alpha + \beta = 1$ будет такой оценкой.

Во-вторых, что менее тривиально, мы имеем равенство $(\varphi, h) = (\varphi^*, f)$, где φ^* — решение уравнения

$$\varphi^*(x) = \int k(y, x) \varphi^*(y) \mu(dy) + h(x) \pmod{\mu}. \quad (4.1.16)$$

Действительно, при любых g_1 и g_2 , для которых определено и конечно скалярное произведение $(g_1, K g_2)$, имеем $(g_1, K g_2) = (g_2, K^* g_1)$, где K^* — оператор с ядром $k(y, x)$.

Отсюда, скалярно умножая первое равенство на φ^* , второе — на φ и вычитая одно из другого, получим:

$$\begin{aligned} \left. \begin{array}{l} \varphi^* \\ \varphi \end{array} \right| & \begin{array}{l} \varphi = K\varphi + f \varphi^* \\ \varphi^* = K^*\varphi^* + h \varphi \end{array} \\ & (\varphi, h) = (\varphi^*, f). \end{aligned}$$

Таким образом, нужный нам функционал (φ, h) может быть вычислен с помощью двойственных оценок:

$$\mathfrak{J}_a^*(\omega_\tau) = \frac{f_0 k_{1,0} \dots k_{\tau,\tau-1} h_\tau}{p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau} \quad (4.1.17)$$

и

$$\mathfrak{J}_c^*(\omega_\tau) = \sum_{l=0}^{\tau} \frac{f_0 k_{1,0} \dots k_{l,l-1} h_l}{p_0^0 p_{0,1} \dots p_{l-1,l}}. \quad (4.1.18)$$

Однако условия согласования, которым должны удовлетворять p^0 и p (и следовательно, множества траекторий ω_τ), отличаются очевидным образом от условий, сформулированных в теоремах 4.1 и 4.2. Так, для $\mathfrak{J}_a^*$ сформулируем их следующим образом:

- (1) $p^0(x) > 0$ для тех x , для которых $f(x) \neq 0 \pmod{\mu}$,
 - (2) $p(x, y) > 0$ для тех x и y , для которых $k(y, x) \neq 0 \pmod{\mu^2}$,
 - (3) $g(x) > 0$ для тех x , для которых $h(x) > 0 \pmod{\mu}$.
- (4.1.19)

Для оценки (4.1.18) должны выполняться первые два условия.

Далее следуют замечания относительно использования оценок при вычислениях.

- Очевидно, что оценки со звездочкой (4.1.17) и (4.1.18) предпочтительнее оценок (4.1.12) и (4.1.15) при вычислении многих функционалов (φ, h) . Первые дают возможность при моделировании одной траектории найти оценки многих функционалов. Вторые для этого не приспособлены: их использование потребует, по крайней мере, дополнительных исследований.
- Также легко заметить, что вычисления существенно упрощаются, если для оценок (4.1.12) и (4.1.14) $p(x, y)$ совпадает с $|k(x, y)|$ или отличается от него на множитель простой структуры. Для оценок (4.1.17) и (4.1.18) $p(x, y)$ следует выбирать близким к $|k(y, x)|$. То же относительно h и p^0 , f и g в первом случае и f и p^0 , h и g во втором.

Примеры. (1) Покажем, в частности, важность учета условий согласования при выборе оценки.

Рассмотрим интегральное уравнение

$$\begin{aligned} \varphi(X) = \frac{1}{(2\pi)^{3/2}(\det C)^{1/2}} \int_D \exp\left(-\frac{1}{2}(Y - X)C^{-1}(Y - X)^T\right) \varphi(Y) dY + \\ + \frac{1}{1 + \sum_{i=1}^s x_i^2}, \end{aligned} \quad (4.1.20)$$

$X = (x_1, \dots, x_s)$, $Y = (y_1, \dots, y_s)$, $D \subset \mathbf{R}^3$ — ограниченная область, $C = \|c_{i,j}\|_{i,j=1}^s$ — положительно определенная матрица. В данном случае интегрирование ведется по σ -конечной мере Лебега. Как мы отмечали ранее, предположение о том, что μ — вероятностная мера, имеет технический характер. Достаточно предположения о ее σ -конечности.

При достаточно большом s (например, $s \geq 20$) задача вычисления (φ, h) для φ , удовлетворяющего (4.1.20), является довольно сложной, а при $s \sim 100$ — практически неразрешимой с помощью детерминированных методов.

Применение метода Монте-Карло в ранее описанной форме приводит к простому алгоритму. В качестве $p(x \rightarrow y)$ целесообразно выбрать многомерную нормальную плотность со средним X и матрицей ковариаций C .

Применение оценки (4.1.12) при этом требует вычисления

$$g(x) = 1 - \frac{1}{(2\pi)^{3/2}(\det C)^{-1/2}} \int_D \exp\left(-\frac{1}{2}(Y - X)C^{-1}(Y - X)^T\right) dY,$$

что может быть далеко не простой задачей, в особенности для сложной области D .

Что касается оценки (4.1.14), то при ее использовании достаточно просто проверять факт выхода X_k за границу D после очередного перехода.

(2) Пусть, как и ранее, D — область в $\mathbf{R}^3$ с границей Γ . Границу Γ предполагаем достаточно гладкой и рассматриваем задачу о вычислении значений функции $u(x)$, $x \in D$, гармонической в D , то есть удовлетворяющей уравнению

$$\Delta u = 0, \quad u|_{\Gamma} = \varphi(X)|_{X \in \Gamma}.$$

Известная теорема о среднем утверждает, что для всякого шара K_R радиуса R с центром в точке X , принадлежащего D , выполняется равенство

$$u(x) = \frac{1}{|K_R|} \int_{|X-Y|=R} U(Y) dY = \frac{1}{|K_R|} \int \delta(|X-Y|=R) \varphi(Y) dY,$$

где $|K_R|$ — площадь поверхности шара K_R , а $\delta(|X-Y|=R)$ — равномерное распределение на поверхности K_R .

Поскольку u известно на границе, можно формально записать уравнение 2-го рода

$$u(x) = \frac{1}{|K_R|} \int_D \delta(|X-Y|=R) \varphi(Y) dY + \varphi(X) \delta(X \in \Gamma), \quad (4.1.21)$$

где $\delta(X \in \Gamma)$ — единичный заряд в точке X поверхности Γ .

Согласно общей схеме, случайный процесс для вычисления $u(x_0)$, $x_0 \in D$, приводящий к простейшей оценке ($k(x, y) = p(x \rightarrow y)$),

должен описываться следующим образом. Траектория частицы начинается в точке x_0 . Затем эта частица переходит в точку, равномерно распределенную на поверхности шара K_{R_0} с центром в точке x_0 . R_0 может быть выбрано произвольным, но таким, чтобы $K_{R_0} \subset D$. Процесс повторяется, и если частица находилась в точке x_i , то ее следующее местонахождение x_{i+1} есть точка, равномерно распределенная на поверхности шара $K_{R_{i+1}}$ с центром в точке x_i , причем $K_{R_i} \subset D$.

Если частица выйдет на границу области, то процесс оборвется. (Обратите внимание на условия согласования для оценки (4.1.12), которые здесь следует использовать. Граница должна служить поглощающим состоянием.) Однако вероятность выйти точно в граничную точку — точку касания со сферой, равна нулю. Процесс оказывается нереализуемым в вычислительном смысле. Следует построить ε -окрестность границы, распространить граничные условия на эту окрестность и обрывать процесс при попадании частицы в нее. Можно считать, что внутри ε -окрестности функция $u(x)$ постоянна вдоль отрезка внутренней нормали, ей принадлежащей.

Можно показать, что дополнительная погрешность, вызванная такой процедурой, имеет порядок ε , то есть оценка (4.1.12) в этом случае будет ε -смещенной оценкой $u(x_0)$ [43]. $\square$

Мы привели лишь формальное описание процесса “блуждания по сферам” как пример процесса, к которому могут привести условия согласования. Подробное исследование этого и родственных процессов содержится в доступной литературе (например, [43], [45]).

4.2. Вторые моменты, понижение дисперсии

Выражения для вторых моментов оценок, полезные для многих целей, могут быть получены путем прямых выкладок.

Для оценки (4.1.12) имеем

$$\begin{aligned} \mathbb{E} \mathfrak{I}_a^2(\omega_\tau) &= \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \\ &\dots \int \mu(dx_\tau) \frac{h_0^2 k_{0,1}^2 \dots k_{\tau-1,\tau}^2 f_\tau^2}{(p_0^0)^2 p_{0,1}^2 \dots p_{\tau-1,\tau}^2 g_\tau^2} p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau = \\ &= \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_\tau) \frac{h_0^2}{p_0^0} \cdot \frac{k_{0,1}^2}{p_{0,1}} \dots \frac{k_{\tau-1,\tau}^2}{p_{\tau-1,\tau}} \cdot \frac{f_\tau^2}{g_\tau}. \end{aligned}$$

Конечность этой суммы есть условие конечности второго момента. Если он конечен, то

$$D\mathfrak{J}_a^2(\omega_\tau) = \left(\frac{h^2}{p^0}, \psi_a \right) - (\varphi, h)^2, \quad (4.2.1)$$

где $\psi_a(x)$ — итерационное решение уравнения

$$\psi_a(x) = \int \frac{k^2(x, y)}{p(x \rightarrow y)} \psi_a(y) \mu(dy) + \frac{f^2(x)}{g(x)} \pmod{\mu}. \quad (4.2.2)$$

Аналогично

$$D\mathfrak{J}_a^*(\omega_\tau) = \left(\frac{f^2}{p^0}, \psi_a^* \right) - (\varphi^*, f)^2, \quad (4.2.3)$$

где ψ_a^* — решение уравнения

$$\psi_a^*(x) = \int \frac{k^2(y, x)}{p(x \rightarrow y)} \psi_a^*(y) \mu(dy) + \frac{h^2(x)}{g(x)}. \quad (4.2.4)$$

Иногда бывают полезны выражения, использующие двойственные к (4.2.2) и (4.2.4) уравнения.

Сложнее вычисляются вторые моменты оценок по столкновениям. Мы проведем вычисления для оценки $\mathfrak{J}_c^*(\omega_\tau)$. Обозначим

$$Q_l = \frac{f_0 k_{1,0} \dots k_{l,l-1}}{p_0^0 p_0 \dots p_{l-1,l}}.$$

Имеем

$$\begin{aligned} \mathfrak{J}_c^*(\omega_\tau) &= \sum_{l=0}^{\tau} Q_l h_l, \\ E(\mathfrak{J}_c^*)^2 &= \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_\tau) \left(\sum_{l=0}^{\tau} Q_l h_l \right)^2 = \\ &= \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_\tau) \sum_{l=0}^{\tau} \sum_{j=0}^{\tau} Q_l h_l Q_j h_j. \end{aligned}$$

Предполагая конечность второго момента и абсолютную сходимость рядов (4.1.4) и (4.1.10), получим

$$\begin{aligned} E(\mathfrak{J}_c^*)^2 &= \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_\tau) \sum_{l=0}^{\tau} (Q_l h_l)^2 p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau + \\ &+ 2 \sum_{\tau=1}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_\tau) \sum_{l=1}^{\tau} \sum_{j=0}^{l-1} Q_l Q_j \cdot h_l h_j p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau. \end{aligned}$$

С первой суммой по τ мы можем обойтись в точности так же, как и при доказательстве теоремы 4.2. Тогда

$$S_1 = \sum_{\tau=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_{\tau}) \sum_{l=0}^{\tau} (Q_l h_l)^2 p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_{\tau} = (\psi^*, h^2),$$

где ψ^* есть итерационное решение уравнения

$$\psi^*(x) = \int \frac{k^2(y, x)}{p(x \rightarrow y)} \psi^*(y) \mu(dy) + \frac{f^2(x)}{p^0(x)}. \quad (4.2.5)$$

Для второй суммы имеем

$$\begin{aligned} S_2 &= 2 \sum_{\tau=1}^{\infty} \int \mu(dx_0) \dots \\ &\dots \int \mu(dx_{\tau}) \sum_{l=1}^{\tau} \sum_{j=0}^{l-1} Q_l Q_j h_l h_j p_0^0 \dots p_{j-1,j} p_{j,j+1} \dots p_{l-1,l} \dots p_{\tau-1,\tau} g_{\tau} = \\ &= \sum_{j=0}^{\infty} \int \dots \int \mu(dx_j) \frac{f_0^2}{p_0} \frac{k_{1,0}^2}{p_{0,1}} \dots \frac{k_{j,j-1}^2}{p_{j-1,j}} h_j \cdot \sum_{l=j+1}^{\infty} k_{j+1,j} \dots \\ &\dots k_{l,l-1} h_l \sum_{\tau=l}^{\infty} \int dx_{l+1} \dots \int dx_{\tau} p_{l,l+1} \dots p_{\tau-1,\tau} g_{\tau}, \end{aligned}$$

полагая $p_{-1,0} \equiv 1$. Мы видим, что это выражение представляет собой произведение трех сомножителей. Последний после изменения нумерации переменных интегрирования на основании (4.1.10) оказывается равным единице. Второй после аналогичного изменения нумерации оказывается равным $\varphi^*(x_j) - h(x_j)$. Имеем $S_2 = 2(\psi^* h, \varphi - h)$. И следовательно,

$$(E\mathfrak{J}_c^*)^2 = 2(\psi^*, h(\varphi^* - h)) + (\psi^*, h^2) = (\psi^*, h(2\varphi^* - h)).$$

Таким образом, справедлива следующая теорема [35], [50].

Теорема 4.3. В предположениях теоремы 4.2 справедливо равенство

$$D\mathfrak{J}_c^* = (\psi^*, h(2\varphi^* - h)) - (h, \varphi)^2, \quad (4.2.6)$$

где ψ^* — итерационное решение уравнения (4.2.5). $\square$

В литературе [35], [81] описываются многие другие несмещенные оценки и вычисляется их дисперсия. Техника доказательства соответствующих утверждений в общих чертах та же, что и при доказательстве приведенных выше теорем.

4.3. Методы уменьшения дисперсии

Как уже отмечалось (см. п. 2.2), развитые в предыдущих разделах этой главы методы оценивания функционалов от решения уравнений 2-го рода используют их представления в виде интегралов по вероятностной мере, индуцированной некоторой цепью Маркова (интегралов по траекториям). Следовательно, необходимо применять общие методы уменьшения дисперсии, что не исключает использования специальных методов, учитывающих конкретные особенности задачи.

Остановимся сначала на приемах, основанных на общих методах, которые описаны в гл. 3.

4.3.1. Метод существенной выборки

Рассмотрим оценку $\mathfrak{J}_a(\omega_\tau)$. Как мы видели, с помощью этой оценки для каждой цепи Маркова, удовлетворяющей условиям согласования, имеем

$$(h, \varphi) = E\mathfrak{J}_a(\omega_\tau) = \int \mathfrak{J}_a(\omega_\tau) P(d\omega_\tau), \quad (4.3.1)$$

где P — мера на траекториях марковской цепи, такая что плотность распределения траектории $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_\tau$ по отношению к мере $\mu^{\tau+1}$ есть $p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau$. Если мы хотим ввести другую меру — $Q(d\omega_\tau)$, такую что существует производная Радона–Никодима dP/dQ , то P должна быть абсолютно непрерывна по отношению к Q . Мера Q должна иметь ту же структуру, что и P . Мы предполагаем ее такой, что траектория ω_τ имеет плотность $q_\tau(x_0, \dots, x_\tau)$ по отношению к $\mu^{\tau+1}$, и тогда q_τ должна быть больше нуля для тех ω_τ , для которых $p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau > 0$.

В соответствии с теоремой 3.1 имеем для оптимальной плотности

$$q_{\tau, \text{опт}}(\omega_\tau) = C |\mathfrak{J}_a(\omega_\tau)| p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau,$$

или

$$q_{\tau, \text{опт}}(\omega_\tau) = C |h_0 k_{0,1} \dots k_{\tau-1,\tau} f_\tau|, \quad (4.3.2)$$

где C — константа нормировки,

$$C^{-1} = \sum_{\tau=0}^{\infty} \int d\mu^{\tau+1} |h_0 k_{0,1} \dots k_{\tau-1,\tau} f_\tau| = (|h|, \bar{\varphi}),$$

а $\bar{\varphi}$ удовлетворяет уравнению

$$\bar{\varphi}(x) = \int \mu(dy) |k(x, y)| \bar{\varphi}(y) + |f(x)|. \quad (4.3.3)$$

Далее мы должны заметить, что необязательно существует цепь Маркова, индуцирующая меру Q . Для существования такой цепи требуется выполнение равенств

$$q_{\tau, \text{опт}}(\omega_\tau) = q_0^0 q_{0,1} \dots q_{\tau-1, \tau} r_\tau,$$

где q^0 является плотностью по отношению к мере μ , а $q = q(x \rightarrow y)$ — неотрицательная $(\text{mod } \mu)$ функция, удовлетворяющая условию

$$\int q(x \rightarrow y) \mu(dy) = 1 - r(x), \quad 0 \leq r(x) \leq 1 \quad (\text{mod } \mu). \quad (4.3.4)$$

Анализируя (4.2.8), мы видим, что $|h|$ не обязательно должна быть плотностью, а для $|k(x, y)|$ не обязательно выполняется равенство

$$\int |k(x, y)| \mu(dy) = 1 - |f(x)|$$

и, конечно же, не обязательно $0 \leq |f(x)| \leq 1 \pmod{\mu}$. Тем не менее в рассматриваемом случае всегда можно указать цепь Маркова, индуцирующую оптимальную в смысле теоремы 3.1 меру. Для этой цели нужно использовать уравнение (4.3.3) и представить $q_{\tau, \text{опт}}$ в виде

$$q_{\tau, \text{опт}}(\omega_\tau) = \frac{|h_0| \bar{\varphi}_0}{(|h|, \bar{\varphi})} \cdot \frac{|k_{0,1}| \bar{\varphi}_1}{\bar{\varphi}_0} \dots \frac{|k_{\tau-1, \tau}| \bar{\varphi}_\tau}{\bar{\varphi}_{\tau-1}} \cdot \frac{|f_\tau|}{\bar{\varphi}_\tau}. \quad (4.3.5)$$

Легко видеть из (4.2.3), что

(1) $\bar{\varphi}_0 > 0$, если $|k_{0,1}| > 0$; $\bar{\varphi}_\tau > 0$, если $|\bar{f}_\tau| > 0$ (операция деления законна),

(2) $\frac{|h_0| \bar{\varphi}_0}{(|h|, \bar{\varphi})}$ является плотностью,

(3) $\int \frac{|k(x, y)| \bar{\varphi}(y)}{\bar{\varphi}(x)} \mu(dx) = 1 - \frac{|f(x)|}{\bar{\varphi}(x)}, \quad 0 \leq \frac{|f(x)|}{\bar{\varphi}(x)} \leq 1;$

все равенства и неравенства следует понимать $(\text{mod } \mu)$.

Из этих рассуждений следует теорема.

Теорема 4.4. Существует цепь Маркова $(q^0(x), q(x \rightarrow y))$, такая что при $p^0 = q^0$, $p(x \rightarrow y) = q(x \rightarrow y)$ достигается минимум дисперсии оценки $\mathfrak{J}_c(\omega_\tau)$. При этом

$$q^0(x) = \frac{|h(x)| \bar{\varphi}(x)}{(|h|, \bar{\varphi})}, \quad q(x \rightarrow y) = \frac{|k(x, y)| \bar{\varphi}(y)}{\bar{\varphi}(x)}, \quad (4.3.6)$$

а $\bar{\varphi}$ есть решение уравнения (4.3.3). Значение дисперсии, соответствующее оптимальной цепи Маркова, есть

$$D_{\text{опт}}\mathfrak{J}_a(\omega_\tau) = (|h|, \bar{\varphi})^2 - (h, \varphi)^2. \quad (4.3.7)$$

□

Доказательство нами проведено ранее. Исключение представляет равенство (4.3.7), которое может быть получено из (4.2.1) путем элементарных выкладок.

Следствие. Если $h \geq 0$, $f \geq 0$, $k \geq 0 \pmod{\mu}$ и $\pmod{\mu^2}$ соответственно, то

$$D_{\text{опт}}\mathfrak{J}_a(\omega_\tau) = 0. \quad (4.3.8)$$

□

Мы видим, что, как и в случае теоремы 3.1, оптимальный результат достигается, если известно решение $\bar{\varphi}$ уравнения (4.3.3). Практическое значение имеет использование априорных сведений о $\bar{\varphi}$.

Исследование дисперсии оценки $\mathfrak{J}_a(\omega_\tau)$ проводится аналогично предыдущему. В этом случае оптимальная цепь Маркова определяется следующим образом:

$$q^{0*}(x) = \frac{|f(x)|\bar{\varphi}^*(x)}{(|f|, \bar{\varphi}^*)}, \quad q^*(x \rightarrow y) = \frac{|k(y, x)|\bar{\varphi}^*(y)}{\bar{\varphi}^*(x)}, \quad (4.3.9)$$

а $\bar{\varphi}^*$ есть решение уравнения

$$\bar{\varphi}^*(x) = \int |k(y, x)|\bar{\varphi}^*(y)\mu(dx) + |h(x)|. \quad (4.3.10)$$

В последнем случае для уменьшения дисперсии используется информация о $\bar{\varphi}^*$ (двойственность), что бывает полезным, особенно если $\bar{\varphi}^*$ имеет некоторый физический смысл.

Не представляет труда формально указать оптимальную меру Q в случае оценок по столкновениям $\mathfrak{J}_c(\omega_\tau)$ и $\mathfrak{J}_c^*(\omega_\tau)$, однако цепи Маркова, соответствующей этой мере, не существует.

Если с самого начала ограничиться мерами, индуцируемыми цепями Маркова, то можно решить задачу об определении оптимальной марковской цепи, но ее решение не вытекает непосредственно из теоремы 3.1, — требуется развитие специального аппарата (см. [58]). Известно, что для различных оценок полезно выбирать переходную плотность цепи Маркова так, чтобы $|k(x, y)|$ (соответственно $|k(y, x)|$) входил в нее в качестве сомножителя.

4.3.2. Понижение порядка интегрирования (векторные оценки)

Если ряд Неймана, с помощью которого представлен функционал

$$(h, \varphi) = \sum_{l=0}^{\infty} \int \mu(dx_0) \dots \int \mu(dx_l) h_0 k_{0,1} \dots k_{l-1,l} f_l, \quad (4.3.11)$$

может быть преобразован путем выполнения аналитического или численного интегрирования по части переменных, то нередко таким образом удастся построить оценки с меньшей дисперсией. Аналогичный прием рассмотрен в главе 3 применительно к общему случаю интеграла по вероятностной мере. Частные особенности суммы (4.2.7) позволяют иметь дело со специальными вычислительными методами.

Отметим прежде всего, что прием понижения порядка интегрирования может применяться к отдельным членам суммы (зачастую удастся вычислить аналитически первые или несколько первых слагаемых ряда). Поскольку переменные x_i обычно векторы $x_i = (x_i^1, \dots, x_i^s)$, то также нередко можно интегрировать все слагаемые по части компонент вектора x_i . Этот случай представляет особый интерес, и мы остановимся на нем подробно.

Пусть $x = (y, z)$ и $\mu(dx) = \mu_2(dy) \mu_1(dz)$. Рассмотрим оценку $\mathfrak{J}_a(\omega_\tau)$ для вычисления (h, φ) :

$$\begin{aligned} \mathfrak{J}_a(\omega_\tau) &= \frac{h_0 k_{0,1} \dots k_{\tau-1,\tau} f_\tau}{p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau}, \\ k_{i,j} &= k(y_i, z_i; y_j, z_j), \quad p_{i,j} = p(y_i, z_i \rightarrow y_j, z_j), \\ h_0 &= h(y_0, z_0), \quad p_0^0 = p^0(y_0, z_0), \\ f_i &= f(y_i, z_i), \quad g_\tau = g(y_\tau, z_\tau); \\ \mathfrak{X} &= \mathcal{Y} \times \mathcal{Z}, \quad y_i \in \mathcal{Y}, \quad z_i \in \mathcal{Z}, \quad i = 0, \dots, \tau. \end{aligned}$$

Предположим, что при любом τ интеграл от $h_0 k_{0,1} \dots k_{\tau-1,\tau} f_\tau$ по $\mu(dz_0) \dots \mu(dz_\tau)$ может быть вычислен точно. С учетом структуры оценки это означает, что при переходе от точки с номером j к $j+1$ вместо умножения на значения ядра $k_{j,j+1}$ числитель оценки $\mathfrak{J}_a(\omega_\tau)$ подвергается действию оператора $L_{j,j+1}$, осуществляющего умножение на $k_{j,j+1}$ и интегрирование по мере $\mu_1(dz_{j+1})$ так, что числитель новой оценки будет иметь вид

$$\int \mu_1(dz_0) h_0 L_{01} \dots L_{\tau-1,\tau}.$$

Если теперь мы выберем цепь Маркова таким образом, чтобы она не зависела от переменных из $\mathcal{Z}$, то есть

$$\bar{p}^0(y, z) = \bar{p}^0(y), \quad p(y, z \rightarrow y', z') = \bar{p}(y \rightarrow y'),$$

то оценка

$$\bar{\mathfrak{J}}_a(\omega_\tau) = \frac{\int \mu_1(dz_0)(h_0 L_{0,1} \dots L_{\tau-1,\tau} f_\tau)}{\bar{p}_0^0 \bar{p}_{0,1} \dots \bar{p}_{\tau-1,\tau} \bar{g}_\tau}, \quad L_{-1,0} \equiv 1 \quad (4.3.12)$$

будет несмещенной при выполнении следующих условий:

- (1) $\bar{p}_0^0 \bar{p}_{0,1} \dots \bar{p}_{\tau-1,\tau} \bar{g}_\tau > 0$ для тех $(\bmod \mu)_2^{\tau+1} y_0, \dots, y_\tau$, для которых также $(\bmod \mu)_2^{\tau+1}$ отлично от нуля выражение

$$\int \mu_1^\tau(dz_0) h_0 L_{0,1} \dots L_{\tau-1,\tau} f_\tau;$$

- (2) конечна сумма ряда

$$\sum_{\tau=0}^{\infty} \int \mu_2(dy_0) \dots \int \mu_2(dy_\tau) \left| \int \mu_1(dz_0) h_0 L_{0,1} \dots L_{\tau-1,\tau} f_\tau \right|. \quad (4.3.13)$$

Тогда имеет место следующая теорема.

Теорема 4.5. При выполнении условий (1) и (2) имеет место равенство

$$E \bar{\mathfrak{J}}_a(\omega_\tau) = (h, \varphi). \quad \square$$

Доказательство тривиально, и мы предоставляем читателю провести его самостоятельно.

Поскольку в исходной оценке $\mathfrak{J}_a(\omega_\tau)$ удобно выбирать $p(x \rightarrow x')$ близкой к $|k(x \rightarrow x')|$, то в рассматриваемом случае возможен выбор таких $p^0(x)$ и $p(x \rightarrow x')$, что плотность траекторий $p_0^0 p_{0,1} \dots p_{\tau-1,\tau} g_\tau$ также удобно проинтегрировать по переменным $z_0, \dots, z_\tau$. Рассмотрим оценку

$$\bar{\bar{\mathfrak{J}}}_a(\omega_\tau) = \frac{\int \mu_1(dz_0)(h_0 L_{0,1} \dots L_{\tau-1,\tau} f_\tau)}{\int \mu_1(dz_0)(p_0^0 \pi_{0,1} \dots \pi_{\tau-1,\tau} g_\tau)}, \quad (4.3.14)$$

где π — оператор, определяемый равенством

$$\pi_{j,j+1} \psi(z_j) = \int \mu_1(dz_{j+1}) p(y_j, z_j \rightarrow y_{j+1}, z_{j+1}) \psi(z_{j+1}),$$

ψ — ограниченная $(\text{mod } \mu_1)$ функция на $\mathcal{Z}$. Оценка (4.3.14) будет несмещенной оценкой (h, φ) при выполнении соответствующих условий согласования (4.3.13), если моделируется цепь Маркова (p_0, p) , но используется лишь траектория $y_0, \dots, y_\tau$.

Следующие замечания являются важными.

Замечания. (1) Условие (4.2.20) слабее условия (4.1.4). Оценки $\bar{\mathfrak{J}}_a$ и $\bar{\bar{\mathfrak{J}}}_a$ могут быть интегрируемыми на траекториях марковской цепи (иметь математическое ожидание), когда $\mathfrak{J}_a$ неинтегрируема.

(2) Если $\mathfrak{J}_a$ имеет математическое ожидание, то $\bar{\mathfrak{J}}_a$ и $\bar{\bar{\mathfrak{J}}}_a$ могут служить средством уменьшения вычислительной работы. Решение о целесообразности их использования требует в конкретных случаях отдельных исследований. Отметим только, что при неотрицательных f, k, h дисперсия $\bar{\bar{\mathfrak{J}}}_a$ обращается в нуль на той же марковской цепи, что и $D\mathfrak{J}_a$. $\square$

“Операторные” аналоги оценок $\mathfrak{J}_a^*$, $\mathfrak{J}_c'$ и $\mathfrak{J}_c^*$ строятся аналогичным образом. Особенности, связанные с их использованием, во многом схожи с таковыми при использовании $\bar{\mathfrak{J}}_c$ и $\bar{\bar{\mathfrak{J}}}_c$.

Особенности операторных оценок полезно проследить на примере систем линейных алгебраических уравнений. Имеем

$$\begin{aligned} X &= AX + F, \\ X &= (x_1, \dots, x_n)^T, \quad A = \|a_{i,j}\|_{i,j=1}^n, \quad F = (f_1, \dots, f_n)^T. \end{aligned}$$

Разобьем множество натуральных чисел $1, 2, \dots, n$ на s непустых подмножеств $N_1 = (n_0 + 1, \dots, n_1)$, $N_2 = (n_1 + 1, \dots, n_2)$, $\dots$, $N_s = (n_{s-1} + 1, \dots, n_s)$, $n_0 = 0$, $n_1 > n_2 > \dots > n_s$. В соответствии с этим разбиением представим векторы X и F в виде $X = (X_1, \dots, X_s)$ и $F = (F_1, \dots, F_s)$, где $X_{j+1} = (x_{n_j+1}, \dots, x_{n_{j+1}})$, $F_{j+1} = (f_{n_j+1}, \dots, f_{n_{j+1}})$, $j = 0, \dots, s-1$, а матрицу A в виде $A = \|A_{k,l}\|_{k,l=1}^s$, где $A_{k,l}$ — прямоугольные подматрицы $A_{k,l} = \|a_{i,j}\|_{i \in N_k, j \in N_l}$. Аналогично, пусть $H = (H_1, \dots, H_s)$, где $H_{l+1} = (h_{n_l+1}, \dots, h_{n_{l+1}})$. Оценка $\bar{\mathfrak{J}}_a$, соответствующая этому разбиению, есть

$$\bar{\mathfrak{J}}_a(\omega_\tau) = \frac{\sum_{l=1}^s H_l A_{i_0, i_1} \dots A_{i_{\tau-1}, \tau} F_\tau}{\bar{p}_{i_0}^0 \bar{p}_{i_0, i_1} \dots \bar{p}_{i_{\tau-1}, \tau} \bar{g}_\tau}, \quad A_{-1,0} \equiv I,$$

где I — единичная подматрица.

Соответствующий алгоритм может рассматриваться как комбинация метода Монте-Карло и итерационных методов. Он имеет большую трудоемкость по сравнению с алгоритмом, использующим $\mathfrak{J}_a$. Учтите

далее замечания, сделанные выше. Очевидны также преимущества, которые могут предоставить операторные оценки при вычислениях на параллельно-векторных вычислительных структурах.

Другое важное свойство операторных оценок вытекает из решения систем интегральных уравнений. Так, в простейшем случае система двух уравнений

$$\varphi_1(x) = \int k_{1,1}(x, y)\varphi_1(y)\mu(dy) + \int k_{1,2}(x, y)\varphi_2(y) + f_1(x) \pmod{\mu}$$

$$\varphi_2(x) = \int k_{2,1}(x, y)\varphi_1(y)\mu(dy) + \int k_{2,2}(x, y)\varphi_2(y) + f_2(x) \pmod{\mu}$$

может рассматриваться как одно уравнение

$$\varphi(j, x) = \int k(j, x; l, y)\varphi(l, y)\mu(dl, dy) + f(j, x), \quad \mu(dl, dy) = \mu(dy) \otimes \mu_1(dl),$$

где $\mu_1(dl)$ — мера с равными весами на множестве $\{1, 2\}$. Индекс l играет роль переменной z в обсуждении общего случая в начале п. 4.3.2.

4.3.3. Ветвящиеся траектории

Одним из универсальных методов уменьшения дисперсии оценок в задаче суммирования ряда Неймана является использование ветвящихся цепей Маркова. Суть метода проще всего пояснить следующим образом. Пусть для вычисления суммы ряда (4.1.7) выбрана одна из рассмотренных оценок и используется цепь Маркова (p^0, p) , удовлетворяющая соответствующим условиям согласования. В качестве одного испытания будем рассматривать множество M траекторий, таких что начальная часть (от x_0 до x_m включительно) у них совпадает, а $x_i^{(j)}$ ($j = 1, \dots, M$) моделируется независимо при $i \geq m$. Считаем $x_m^{(1)} = x_m^{(2)} = \dots = x_m^{(M)}$. Имеем общую часть $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_\tau$ при $\tau \leq m$ и

$$\begin{array}{c} \nearrow x_{m+1}^{(1)} \rightarrow \dots \rightarrow x_{\tau(1)}^{(1)}, \\ x_0 \rightarrow \dots \rightarrow x_m \longrightarrow \dots, \\ \searrow x_{m+1}^{(M)} \rightarrow \dots \rightarrow x_{\tau(M)}^{(M)}, \end{array} \quad (4.3.15)$$

если $\tau > m$. Пусть $\sum_{j=1}^M \lambda_j = 1$.

Оценка

$$\mathfrak{J}_a^b = \begin{cases} \sum_{j=1}^M \lambda_j \frac{h_0 k_{0,1} \dots k_{m_j, m_j+1}^{(j)} \dots k_{\tau(j)-1, \tau(j)}^{(j)} f_{\tau(j)}^{(j)}}{p_0^0 p_{0,1} \dots p_{m_j, m_j+1}^{(j)} \dots p_{\tau(j)-1, \tau(j)}^{(j)} g_{\tau(j)}^{(j)}}, & \tau(j) \geq m, \\ \frac{h_0 k_{0,1} \dots k_{\tau-1, \tau} f_{\tau}}{p_0^0 p_{0,1} \dots p_{\tau-1, \tau} g_{\tau}}, & \tau(j) = \tau < m, \end{cases} \quad (4.3.16)$$

где $k_{p,q}^{(j)} = k(x_p^{(j)}, x_q^{(j)})$, $p_{p,q}^{(j)} = p(x_p^{(j)} \rightarrow x_q^{(j)})$, $f_q^{(j)} = f(x_q^{(j)})$, $g_q^{(j)} = g(x_q^{(j)})$, является несмещенной. Здесь m_j имеет, вообще говоря, свое значение для каждой траектории и может зависеть от $x_i^{(j)}$. Для данной траектории (4.3.15) точка x_m называется точкой ветвления, а λ_j — статистическим весом j -й ветви. Как правило, выбираем $\lambda_j > 0$. Очевидно, что конструкция обобщается на случай, когда любая из j -х траекторий в свою очередь имеет точку ветвления. Наличие ветвления, значения m , x_m и λ_j влияют на величину дисперсии оценок. (Аналоги $\mathfrak{J}_a^*$, $\mathfrak{J}_c$ и $\mathfrak{J}_c^*$ для ветвящихся траекторий строятся очевидным образом.) Большое количество параметров сильно усложняет задачу оптимизации. Обычно используют ветвление в той области фазового пространства траекторий, которая может значительно повлиять на дисперсию оценки. Эта область определяется экспериментально или на основе “физических” соображений. Здесь важную роль играет двойственность φ и φ^* , но это специальная тема, развиваемая в связи с задачами переноса излучения через вещество. Отметим в заключение, что использование расслоенной выборки при вычислении ряда Неймана также приводит к конструкциям, схожим с ветвящимися цепями Маркова.

4.4. Замечания о трудоемкости.

Квазислучайные числа

Прежде всего следует повторить общее положение: уменьшение дисперсии не может служить самоцелью. Целью теоретических исследований, связанных с применением метода Монте-Карло, является

- (1) доказательство существования математического ожидания, несмещенности или асимптотической несмещенности выбранной оценки;
- (2) оценка сверху второго момента, если он существует;

- (3) выбор меры, подлежащей моделированию, и оценки, обеспечивающей наименьшую трудоемкость (наименьшее среднее число операций компьютера). Заметим, что обычно речь идет не об оптимальном, а о более или менее близком к оптимальному выбору. Выбор такого рода нередко осуществляется путем численных экспериментов и существенно зависит от конкретных особенностей задачи.

В отличие от классических детерминированных методов, теория которых ориентирована на исследование алгоритмов решения обширного класса задач, метод Монте-Карло является инструментом решения конкретной задачи. Поясним последнее утверждение примером. Метод существенной выборки указывает оптимальное распределение узлов интегрирования для заданных f и μ . При детерминированной постановке задача строится для класса функций f и может быть решена в очень ограниченном числе случаев. Тем не менее нужны веские основания, чтобы предпочесть метод Монте-Карло детерминированному алгоритму в тех случаях, когда последний удобен и подробно исследован.

4.4.1. Сравнительная трудоемкость при решении систем линейных алгебраических уравнений

Представляет интерес сравнение детерминированных алгоритмов и алгоритмов метода Монте-Карло в достаточно простом случае. Допустим в предположениях п. 4.1 вычисляется решение системы линейных алгебраических уравнений

$$X = AX + F, \quad \lambda_1(|A|) < 1.$$

Для определенности будем использовать оценку $\mathfrak{J}_c^*$, которая позволяет на одной траектории цепи Маркова оценивать все компоненты решения X и сравнивать ее трудоемкость с трудоемкостью метода простых итераций:

$$X_m = AX_{m-1} + F, \quad X_0 = F.$$

Пусть M — число итераций, необходимых для получения решения с точностью ϵ (в некоторой векторной норме):

$$X_M = (I + A + \dots + A^M)F. \quad (4.4.1)$$

С помощью случайного вектора можно оценить компоненты вектора $X_M = (x_1^{(M)}, \dots, x_n^{(M)})$:

$$[\mathfrak{J}_c^*]_M(l) = \sum_{j=0}^{\tau_M} \frac{f_{i_0} a_{i_1, i_0} \dots a_{i_j, i_{j-1}} h_{i_j, j}}{p_{i_0} p_{i_0, i_1} \dots p_{i_{j-1}, i_j}}, \quad (4.4.2)$$

$$\tau_M = \begin{cases} \tau, & \tau < M, \\ M, & \tau \geq M, \end{cases} \quad a_{0, -1} \equiv 1, \quad h_{i_j, j} = \begin{cases} 1, & i_j = l, \\ 0, & i_j \neq l. \end{cases}$$

Полагая

$$Q_{i_j} = \frac{f_{i_0} a_{i_1, i_0} \dots a_{i_j, i_{j-1}}}{p_{i_0} p_{i_0, i_1} \dots p_{i_{j-1}, i_j}}, \quad i_j \leq \tau_M,$$

можно указать оценку для каждого $x_l^{(M)}$ на одной моделируемой траектории. А именно

$$\hat{x}_l^{(M)} = \begin{cases} \sum_{i_j=l} Q_{i_j}, \\ 0, & \text{если нет } i_j, \text{ равных } l. \end{cases} \quad (4.4.3)$$

Очевидно, $E\hat{x}_l^{(M)} = x_l^{(M)}$.

Рассмотрим случай, когда в каждой строке матрицы имеется ровно r , $r \leq n$, элементов. Для тех случаев, когда среднее число таких элементов примерно есть n , легко сделать качественные выводы на основе дальнейших рассуждений. Вычислим число арифметических операций.

- В случае метода простых итераций для вычисления X_M нужно умножить M раз матрицу A на вектор и M раз сложить векторы. На это требуется $R_n = M(n(2r - 1 + 1)) = 2Mnr$ операций.
- Для вычисления вектора с компонентами $[\mathfrak{J}_c^*]_M(l)$ требуется не более M раз моделировать дискретное распределение (строку переходной матрицы $\|p_{i,j}\|$), осуществить не более M умножений и делений и не более M сложений (добавление Q_{i_j} к ранее полученным оценкам $\hat{x}_l^{(M)}$). В силу условий согласования матрица $\|p_{i,j}\|$ может иметь в строке r ненулевых элементов. Моделирование такого распределения потребует $\log_2(r+1)$ операций (единица добавляется на проверку обрыва траектории). Таким образом, вычисление оценки $[\mathfrak{J}_c^*]_M(l)$ потребует не более $M(\log_2(r+1)+3)$ операций. Отметим, что вычисляются одновременно значения оценки для $l = 1, \dots, n$.

Эта процедура должна повторяться N раз, и в конце вычисляется среднее арифметическое всех n полученных сумм. Следовательно, необходимое число операций:

$$S_n \leq NM(\log_2(r+1) + 3) + n.$$

Получаем

$$\begin{aligned} \frac{S_n}{R_n} &\leq \frac{NM(\log_2(r+1) + 3) + n}{2Mnr} = \\ &= \frac{N \log_2(r+1) + 3}{2nr} + \frac{1}{2Mr} \sim \frac{N \log_2 r}{2nr} + \frac{1}{2Mr}. \end{aligned} \quad (4.4.4)$$

Как известно из общей теории итерационных методов, $M \sim \ln \varepsilon / \ln |\lambda_1|$, где ε — норма погрешности решения, а λ — наибольшее по модулю собственное число A , $|\lambda_1| < 1$.

С другой стороны, при заданном уровне надежности $N \sim \sigma^2 / \varepsilon^2$, где σ^2 — максимальная дисперсия оценок компонент вектора X_M . Асимптотически по ε имеем

$$\frac{S_n}{R_n} \sim \frac{N \log_2 r}{2nr} + \frac{1}{2Mr} = \frac{\sigma^2 \log_2 r}{\varepsilon^2 2nr} + \frac{\ln |\lambda_1|}{2r \ln \varepsilon}. \quad (4.4.5)$$

Так, при $r = \text{const}$, не зависящем от n ,

$$\frac{S_n}{R_n} \sim C_1 \frac{1}{\varepsilon^2 n} + C_2 \frac{1}{\ln \varepsilon}, \quad (4.4.6)$$

и если $n \rightarrow \infty$, $\varepsilon \rightarrow 0$ так, что $\varepsilon^2 n \rightarrow \infty$, то метод Монте-Карло оказывается эффективнее (с заданным уровнем надежности), чем метод простых итераций. Это утверждение тем более справедливо, если r стремится к бесконечности при $n \rightarrow \infty$.

Из сказанного вытекает важный вывод: метод Монте-Карло следует применять в задачах большой размерности.

Это, безусловно, вывод общего характера, не учитывающий многих других обстоятельств.

Метод Монте-Карло

- обладает свойствами естественного параллелизма;
- может оценивать один или несколько функционалов от решения, не вычисляя самого решения;
- эффективен в тех случаях, когда исходные данные заданы в виде случайного процесса (двойная рандомизация).

Однако применительно к системам линейных уравнений

- все сказанное справедливо при условии $|\lambda_1(|A|)| < 1$;
- погрешность носит вероятностный характер;
- необходимая вычислительная работа с уменьшением погрешности возрастает как $1/\varepsilon^2$.

4.4.2. Метод *квази*Монте-Карло

Как мы уже отмечали в п. 3.2, методы *квази*Монте-Карло состоят в применении кубатурной формулы

$$\int_0^1 dx_1 \dots \int_0^1 dx_s f(x_1, \dots, x_s) \approx \frac{1}{N} \sum_{j=1}^N f(x_1^{(s)}(j), \dots, x_s^{(s)}(j)). \quad (4.4.7)$$

Здесь $x_i^{(s)}(j)$ — неслучайные числа из промежутка $(0, 1)$ (то есть они получены не с помощью датчика случайных чисел!). Идея метода *квази*Монте-Карло состоит в том, чтобы использовать $x_i^{(s)}(j)$ в схеме метода Монте-Карло вместо реализации α_l равномерно распределенной на $(0, 1)$ случайной величины. Возвращаясь к материалу главы 1, мы вспоминаем, что α_l можно выбирать в любом порядке. Этого, вообще говоря, нельзя делать с числами $x_i^{(s)}(j)$, хотя во многих схемах *квази*Монте-Карло (например, точки Холтона) при переходе от s к $s + 1$ можно сохранять полученные ранее $x_i^{(s)}(j)$, то есть $x_i^{(s)}(j)$ можно выбирать не зависящими от s , и при увеличении N можно не пересчитывать полученную ранее сумму. Новые узлы добавляются к старым. Эти особенности более свойственны методу Монте-Карло, нежели классической теории кубатурных формул. В последней изменение N или s требует выбора новой системы узлов.

Таким образом, если методом Монте-Карло вычисляется интеграл и подынтегральная функция как функция α_l является функцией ограниченной вариации (в смысле Харди–Краузе), то применение формулы (4.4.6.) может повысить эффективность алгоритма (уменьшить вычислительную работу). При этом о поведении погрешности можно судить по результатам вычислений, полученным при возрастании N .

Если вычисления состоят в моделировании траекторий цепи Маркова и определении соответствующей оценки, то для одной траектории должна использоваться система узлов, принадлежащая одному значению j , то есть $x_1(j), x_2(j), \dots, x_s(j)$. Поскольку траектории могут иметь разную длину, важно, чтобы при возрастании s не нужно было бы изменять полученные ранее $x_i(j)$. Имеется много публикаций, в которых последовательности Холтона и Соболя приводят

к лучшим результатам, нежели стандартный датчик случайных чисел. Однако есть ряд недостаточно исследованных вопросов, связанных с применением метода *квази*Монте-Карло.

Так, известно, что распределения, получаемые в качестве результатов при моделировании естественных явлений, могут иметь искажения вследствие корреляций между элементами квазислучайных последовательностей. Эти искажения бывают особенно заметными, когда применяются методы отбора. Поэтому следует проявлять осторожность при решении задач, использующих моделирование распределений и, в особенности, методы отбора для моделирования распределений. Одна из удобных схем применения метода *квази*Монте-Карло обсуждается в следующей главе.

4.5. Библиографическое дополнение к главе 4

Оценку погрешности решения $\varphi_\varepsilon(x)$ можно получить, используя его представление с помощью функции Грина ^{*)}. В нашем случае имеем

$$\Delta U = -f(x), \quad f(x) = U(x)/\Gamma_\varepsilon.$$

Функция $f(x)$ ограничена и отлична от нуля лишь в ε -окрестности поверхности Γ . Важно отметить, что введение ε -смещенной оценки позволяет пользоваться средствами вычислительной математики, обходя теоретико-вероятностные методы исследования процесса блужданий по сферам.

Существует бесконечно много оценок функционалов (h, φ) , линейных относительно f и h ; об этом см. [35], а также оригинальную работу Хисамутдинова [81]. Обычно ограничиваются случаями простейших в вычислительном отношении оценок.

Многие прикладные задачи, описываемые марковскими цепями, в том числе задачи теории переноса излучения, сводятся к решению интегрального уравнения в форме

$$\varphi(x) = \int k(y, x)\varphi(y)dy + f(x),$$

где $k(y, x)$ имеет смысл переходной (субстохастической) плотности.

В более строгом подходе процесс переноса излучения описывается уравнением относительно мер:

$$d\varphi(x) = \int p(y, dx)d\varphi(y) + df(x).$$

^{*)} См., например: Владимиров В. С. Уравнения математической физики. М., Наука, 1967. С. 342. Формула (17).

Случай уравнения относительно мер подробно рассмотрен в [45]. Он имеет определенные особенности, которые проявляются, прежде всего, при решении двойственной задачи (сопряженного уравнения).

Связь имитационных моделей и интегральных уравнений в случае задач переноса излучений впервые освещена подробно и достаточно строго в работах Фролова и Ченцова [79].

Выражения для дисперсии оценок, использующие прямые и двойственные представления для широкого класса оценок, приведены в книге [35].

Относительно применения теоремы о существенной выборке к вычислению интеграла по траекториям марковской цепи следует заметить, что такой подход оказывается плодотворным лишь в случае оценки по поглощению. Оценка по столкновениям уже не допускает простого рецепта выбора цепи Маркова, минимизирующей дисперсию. Мера на траекториях, обращающая дисперсию в нуль в знакопостоянном случае, не является марковской (не порождается марковской цепью по крайней мере в пространстве состояний, на котором определена φ). Однако можно поставить задачу о выборе цепи Маркова, соответствующей минимуму дисперсии оценок по столкновениям (или других несмещенных оценок). Задача исследована в работах Михайлова [57], [58].

Задачи, возникающие в связи с решением дифференциальных и интегральных уравнений конкретных классов, чрезвычайно разнообразны, поэтому литература, в которой они обсуждаются, весьма обширна. Мы сошлемся на книги [35], [45], [56], [57], [65]. Из работ иностранных авторов интерес представляют [72], [19]. Это далеко не полный перечень наиболее доступных монографий. Самые важные приложения, обсуждаемые в них, — это задачи теории переноса излучения, атмосферной оптики, краевые задачи для уравнений в частных производных и ряд других (см. библиографические дополнения к главам 5 и 6).

Векторные оценки предложены И.М. Соболев. Они исследовались в работах Ермакова [35], Михайлова [57], широко применялись Михайловым при решении прикладных задач [56].

Глава 5

Итерационные процессы. Стохастическая устойчивость

5.1. Разностные схемы и схема Неймана–Улама

Как отмечалось в предыдущей главе, метод Монте-Карло может быть эффективным средством решения систем линейных алгебраических уравнений лишь в случае их достаточно большой размерности. Известно, что системы очень большой размерности нередко возникают при дискретизации многомерных задач математической физики. Случай дискретного анализа простейшего эллиптического уравнения (Пуассона) рассматривался нами ранее.

В дальнейшем мы также рассмотрим простейшие случаи различных типов уравнений, что даст нам возможность сосредоточиться на обсуждении структуры алгоритмов, которая сохраняет общие черты при усложнении задач. Всюду предполагается достаточная гладкость начальных и граничных условий и существование классического (гладкого) решения.

Простейший разностный аналог уравнения теплопроводности

Для уравнения

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2},$$
$$u = u(x, t),$$

в области $0 \leq t \leq T$, $0 \leq x \leq 1$, на сетке с шагом $h = 1/M$ на отрезке $[0, 1]$ и шагом $\Delta t = T/L$ на отрезке $[0, T]$ имеем простейшую

разностную схему

$$\begin{aligned} \frac{u(kh, (l+1)\Delta t) - u(kh, l\Delta t)}{\Delta t} &\approx \\ \approx a^2 \frac{u((k+1)h, l\Delta t) - 2u(kh, l\Delta t) + u((k-1)h, l\Delta t)}{h^2}, \end{aligned}$$

$k = 1, \dots, M-1$; $l = 0, 1, \dots, L-1$. Заменяв приближенное равенство точным, получим систему уравнений для неизвестных $u_{k,l}$, приближенно равных $u(kh, l\Delta t)$.

Если рассматривается первая краевая задача (u задано на границе области), то известными оказываются $u_{k,0}$, $k = 1, \dots, M-1$, а также $u_{0,l}$ и $u_{M,l}$, $l = 0, \dots, L-1$.

Полученная в результате разностная схема называется явной. Если $u_{k,l}$ известны при некотором l , то по формуле

$$\begin{aligned} u_{k,l+1} &= u_{k,l} + \frac{1}{2\mu^2}(u_{k+1,l} - 2u_{k,l} - u_{k-1,l}), \\ \mu^2 &= \frac{h^2}{2\Delta t a^2}, \end{aligned} \quad (5.1.1)$$

легко вычисляются $u_{k,l+1}$. Однако, как известно, в теории разностных схем при $\mu^2 < 1$ такие вычисления оказываются практически невозможными, так как возникает вычислительная неустойчивость, вызываемая накоплением ошибок округления.

При $\mu^2 = 1$ имеем

$$u_{k,l+1} = \frac{1}{2}(u_{k+1,l} + u_{k-1,l}). \quad (5.1.2)$$

Очевидно, что при вычислениях по этой формуле ошибка накапливаться не может, но уже при $\mu^2 = 1/2$ неустойчивость можно наблюдать с помощью так называемой ε -схемы. В этом случае (5.1.1) выглядит следующим образом

$$u_{k,l+1} = u_{k+1,l+1} + u_{k+1,l-1} - u_{k,l}. \quad (5.1.3)$$

Считая все начальные значения нулевыми, кроме одного (например, $u_{\frac{m}{2},0} = \varepsilon$ при четном m), имеем при m достаточно большом, чтобы исключить влияние границы, значения $u_{\frac{m}{2},l}$, приведенные в следующей таблице

l	0	1	2	3	4	5	6
$u_{\frac{m}{2},l}$	ε	$-\varepsilon$	3ε	-7ε	19ε	-51ε	141ε

(5.1.4)

Наблюдается экспоненциальный рост (более чем в два раза) модуля ошибки с ростом l при изменении знака. Это типично для явления неустойчивости. Подобный прием используется для обнаружения неустойчивости сложных разностных схем. В простых случаях, подобных рассматриваемому, возможен теоретический анализ. Для такого анализа в (5.1.1) нужно оценить норму оператора перехода с l -го на $(l+1)$ -й слой.

Далее если формально применить схему Неймана–Улама к системе уравнений (5.1.2), то в соответствии с условиями согласования (см. гл. 4) будем иметь цепь Маркова следующего типа. Для нахождения решения в точке с номером (k_0, l_0) переходим с положительными вероятностями в одну из точек: $(k_0 + 1, l_0 - 1)$ или $(k_0 - 1, l_0 - 1)$, и вообще, если мы находимся во внутренней точке (k', l') , то в следующий момент переходим с положительными вероятностями в одну из точек: $(k' + 1, l' - 1)$ или $(k' - 1, l' - 1)$. В простейшем случае эти вероятности считаем равными, что приводит к удобному алгоритму. Если точка, в которую совершен переход, находится на границе области, то траектория обрывается. Характерной особенностью процесса является то, что длина его траектории не может превышать $N - l_0$.

Очевидным недостатком явной схемы (5.1.1) является необходимость выполнения условия устойчивости, требующего выбора малого шага по времени. Этого недостатка лишена неявная схема

$$u_{k+1,l+1} - (2 + 2\mu^2)u_{k,l+1} + u_{k-1,l+1} = -2\mu^2 u_{k,l}, \quad (5.1.5)$$

которая позволяет при известных $u_{k,l}$, $k = 1, \dots, m-1$, вычислять $u_{k,l+1}$. Для ее анализа при нулевых начальных и граничных условиях можно изучать ненулевые (собственные) решения (5.1.5). Легко проверить, что они имеют вид

$$u_{k,l} = \lambda^l \cdot \sin \frac{\pi r}{m} k, \quad \lambda \neq 0, \quad r = 1, \dots, m-1, \quad (5.1.6)$$

где λ определяется из (5.1.5):

$$\lambda^{l+1} \left(\sin \frac{\pi r}{m} (k+1) + \sin \frac{\pi r}{m} (k-1) - 2(1 + \mu^2) \sin \frac{\pi r}{m} k \right) = -2\mu^2 \lambda^l \sin \frac{\pi r}{m} k.$$

Имеем

$$\lambda \left(2 \cos \frac{\pi r}{m} - 2(1 + \mu^2) \right) \sin \frac{\pi r}{m} k = -2\mu^2 \sin \frac{\pi r}{m} k,$$

или

$$\lambda = \frac{\mu^2}{\mu^2 + \frac{1}{2} \sin^2 \frac{\pi r}{2m}} < 1. \quad (5.1.7)$$

Таким образом, для данной схемы оператор перехода имеет норму меньше единицы при любом соотношении между h и Δt .

Уравнения (5.1.5) могут быть записаны в форме

$$u_{k,l+1} = \frac{1}{2(1+\mu^2)}(u_{k+1,l+1} + u_{k-1,l+1}) + \frac{1}{1+\mu^2}u_{k,l}, \quad (5.1.8)$$

$k = 1, \dots, M-1$; $l = 0, 1, \dots, L-1$. Если U — вектор с компонентами $u_{k,l}$, то при соответствующей их сплошной нумерации матрица системы (5.1.8) будет иметь вид

$$\mathfrak{A} = \begin{pmatrix} A_1 & A_2 & 0 & \dots & 0 \\ 0 & A_1 & A_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & A_1 & A_2 \\ 0 & 0 & \dots & 0 & A_1 \end{pmatrix}, \quad (5.1.9)$$

где A_1, A_2 — матрицы порядка $m-1$,

$$A_1 = \frac{1}{1+\mu^2} \begin{pmatrix} 0 & \frac{1}{2} & 0 & \dots & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \dots & 0 & 0 \\ 0 & 0 & \dots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \dots & 0 & \frac{1}{2} & 0 \end{pmatrix},$$

$$A_2 = \frac{1}{1+\mu^2}I,$$

I — единичная матрица. Легко проверить, что условие $P(|\mathfrak{A}|) < 1$ выполнено и к системе (5.1.8) применима схема Неймана–Улама. Условия согласования определяют характер случайных блужданий. В данном случае можно в качестве переходной матрицы $\mathfrak{P}$ выбрать просто $\mathfrak{A}$. Читателю рекомендуется подробно рассмотреть случайные процессы, согласованные с $\mathfrak{P}$ и $\mathfrak{P}^T$. Эти процессы можно трактовать как процессы с взаимно обратным течением времени.

Отметим, что при фиксированном l систему уравнений (5.1.8) относительно $u_{k,l+1}$ можно решить методом Монте-Карло. Если так поступать при каждом l , то мы получим случайную ошибку метода Монте-Карло для вычисления $u_{k,l+1}$ и ошибку, проистекающую от того, что $u_{l,k}$ также имело случайную ошибку, накопленную в процессе предыдущих вычислений. К анализу поведения этой ошибки мы вернемся позже. Поучительно сравнить этот алгоритм с алгоритмом, реализующим процесс, согласованный с $\mathfrak{P}^T$.

Напомним в очередной раз, что метод Монте-Карло, конечно же, следует применять в многомерном случае, то есть для уравнений

$$\frac{\partial u}{\partial t} = \Delta u, \quad \Delta u = \sum_{i=1}^s \frac{\partial^2 u}{\partial x_i^2},$$

где s по крайней мере больше 2, а также для более сложных, чем Δ , эллиптических операторов в областях сложной формы и т. п. При этом основные особенности алгоритмов те же, что и при $s = 1$.

Рассмотрим теперь, также в простейшем случае, разностные схемы, связанные с волновым уравнением

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \quad (5.1.10)$$

для которого обычно указываются начальные данные

$$u(0, x) = \varphi_1(x), \quad \frac{\partial u}{\partial t}(0, x) = \varphi_2(x)$$

и на концах некоторого отрезка — пусть это будет $[0, 1]$ — граничные условия вида

$$\begin{aligned} \alpha_1(t) \frac{\partial u}{\partial x}(t, 0) + \alpha_2(t) u(t, 0) &= \varphi_3(t), \\ \beta_2(t) \frac{\partial u}{\partial x}(t, 1) + \beta_2(t) u(t, 1) &= \varphi_4(t), \end{aligned}$$

где $\alpha_1, \alpha_2, \beta_1, \beta_2, \varphi_1, \varphi_2, \varphi_3$ и φ_4 — заданные функции.

Используя ту же сетку, что и в случае параболического уравнения, имеем для каждой внутренней точки явную схему:

$$\frac{u_{k,l+1} - 2u_{k,l} + u_{k,l-1}}{(\Delta t)^2} = \frac{u_{k+1,l} - 2u_{k,l} + u_{k-1,l}}{h^2}. \quad (5.1.11)$$

При $\Delta t = h$ из (5.1.11) следует

$$u_{k,l+1} = u_{k+1,l} + u_{k-1,l} - u_{k,l-1}. \quad (5.1.12)$$

Значения $u_{k,l}$ на первых двух слоях по времени ($l = 0, l = 1$) определяются приближенно из начальных условий. Значения на границе исключаются с помощью граничных условий. Ниже рассматривается простейший случай 1-й краевой задачи.

Анализ оператора перехода для схемы (5.1.12) можно осуществить, снова используя представление (5.1.6). Для λ получим квадратное уравнение, имеющее корни комплексные и равные по модулю

единице. Это означает, что вычисления по формулам (5.1.1 – 5.1.12) устойчивы относительно ошибки округления. Тем не менее попытки решить эту систему с помощью случайных блужданий не приводят к успеху. Структура “большой” матрицы такова:

$$\mathfrak{B} = \begin{pmatrix} 0 & B_1 & B_2 & 0 & \dots & 0 \\ 0 & 0 & B_1 & B_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & \dots & & 0 \end{pmatrix}, \quad (5.1.13)$$

$$B_1 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & & & \dots & 1 & 0 \end{pmatrix}, \quad B_2 = -I.$$

Условие $P(|\mathfrak{B}|) < 1$ здесь выполнено, но для всякой цепи Маркова, удовлетворяющей условиям согласования, отношения a_{ij}/p_{ij} будут больше единицы по абсолютной величине. Соответствующие оценки будут менять знак на траекториях. При конечном N — числе шагов по времени — дисперсия оценки для каждого N остается ограниченной, но растет с ростом N экспоненциально. Это следует из того, что решение уравнения с матрицей $\|a_{ij}^2/p_{ij}\|$ в данном случае экспоненциально растет с ростом N .

Мы имеем дело с явлением стохастической неустойчивости алгоритма метода Монте-Карло.

Неявная схема для волнового уравнения в простейшем случае:

$$U^{l+1} = A_1 U^{l+1} + A_2 U^l + A_3 U^{l-1} + \mathcal{F}^l, \quad (5.1.14)$$

где U^l есть вектор неизвестных, относящихся к l -му слою по времени. U^0 и U^1 предполагаются известными. Задачу о стохастической устойчивости, связанной с этой схемой алгоритма метода Монте-Карло, мы рассмотрим в следующих разделах, а сейчас остановимся на проблеме стохастической устойчивости итерационных процессов.

Пусть имеется система линейных алгебраических уравнений

$$X = AX + \mathcal{F} \quad \text{и} \quad |\lambda_1(A)| < 1.$$

Разобьем матрицу A на два слагаемых $A = A_1 + A_2$ и рассмотрим итерационный процесс

$$X_n = A_1 X_n + A_2 X_{n-1} + \mathcal{F}, \quad n = 1, 2, \dots, \quad X_0 = \mathcal{F}, \quad (5.1.15)$$

при этом предположим, что

$$|\lambda_1(|A_1|)| < 1, \quad |\lambda_1((I - A_1)^{-1}A_2)| < 1. \quad (5.1.16)$$

Для каждого n будем находить X_n с помощью метода Монте-Карло (схема Неймана–Улама). Это означает, что строится последовательность Ξ_n несмещенных оценок векторов X_n . Дисперсия Ξ_n , которую мы предполагаем конечной, будет зависеть от числа независимых траекторий N_n ($D\Xi_1 \sim 1/N_n$) и от других факторов — вида оценки и т. п. В первую очередь будет учитываться зависимость от N_n . Последовательность вторых моментов Ξ_n может оставаться ограниченной, но может и неограниченно возрастать с ростом n . В последнем случае алгоритм оказывается стохастически неустойчивым. Задачи о стохастической устойчивости разностных схем и итерационных алгоритмов (5.1.15) очевидным образом тесно связаны. $(I - A_1)^{-1}A_2$ для разностной схемы является оператором перехода. Обозначим его норму через $\rho = |\lambda_1((I - A_1)^{-1}A_2)|$. Условие $\rho < 1$ является необходимым и достаточным условием сходимости итерационного процесса (5.1.15) и достаточным условием устойчивости разностной схемы вида (5.1.15). В разностной схеме вычисления ведутся до некоторого конечного $n = n_0$, и для $\rho = 1$ ошибки округления остаются ограниченными. В этом случае схема устойчива относительно ошибок округления [66].

5.2. Исследование стохастической устойчивости

Пусть векторы X_l , $l = 1, 2, \dots$, вычисляются по формуле

$$X_l = A_1 X_l + A_2 X_{l-1} + \mathcal{F}_l, \quad (5.2.1)$$

где $X_0, \mathcal{F}_l$ — заданные векторы, $A_1 = \|a_{i,j}^1\|_{i,j=1}^n$, $A_2 = \|a_{i,j}^2\|_{i,j=1}^n$ — заданные матрицы*). Полагаем, что при фиксированном l вектор X_l

*) Более общий случай соотношений $X_{l+k} = B_1 X_{l-k-1} + \dots + B_k X_l + \mathcal{F}_l$, $B_r = \|a_{i,j}^r\|_{i,j=1}^n$, $r = 1, \dots, k$, сводится к данному введением в рассмотрение вектора $U_l = (X_{l+k}, \dots, X_{l+k+1})$, если положить, например,

$$A_1 = \begin{pmatrix} B_1 & \dots & B_{k-1} \\ 0 & \dots & 0 \\ 0 & \dots & 0 \\ \cdot & \cdot & \cdot \\ 0 & \dots & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 & \dots & B_k \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

находится с помощью метода Монте-Карло. Для определенности полагаем, что используется схема Неймана–Улама, описанная в гл. 4, $(|\lambda_1(|A_1|)|) < 1$, так что вычисляются несмещенные оценки X_l по N_l траекториям. Считаем, что N_l не зависит от l (то есть имеет место однородность оценок). Таким образом, вычисления в действительности производятся по формулам

$$\Xi = (\widetilde{I - A_1})^{-1} \widetilde{A_2} \Xi_{l-1} + (\widetilde{I - A_1})^{-1} \mathcal{F}_l, \quad (5.2.2)$$

где волна означает, что соответствующий результат обращения матриц, их умножения и умножения на вектор Ξ_{l-1} осуществляется с помощью метода Монте-Карло. Заметим, что случайный оператор $(\widetilde{I - A_1})^{-1} \widetilde{A_2}$ статистически не зависит от Ξ_{l-1} в соответствии со структурой алгоритма. Конструкция этого оператора и вычисление Ξ_{l-1} производится с помощью различных случайных чисел.

Кроме того,

$$\begin{aligned} \mathbb{E}[(\widetilde{I - A_1})^{-1}] &= (I - A_1)^{-1}, \\ \mathbb{E}[(\widetilde{I - A_1})^{-1} \widetilde{A_2}] &= (I - A_1)^{-1} A_2, \\ \mathbb{E} \Xi_l &= X_l. \end{aligned} \quad (5.2.3)$$

При сделанных предположениях изучим поведение ковариационных матриц векторов Ξ_l в зависимости от l .

Обозначим $\Xi_l - X_l = \mathcal{E}_l$ — вектор-столбец случайных погрешностей. Тогда его матрица ковариаций есть $\text{cov} \Xi_l = \mathbb{E}(\mathcal{E}_l \mathcal{E}_l^T)$. Получим выражение $\text{cov} \Xi_l$ через $\text{cov} \Xi_{l-1}$ в удобной форме.

Из (5.2.2) имеем

$$X_l + \mathcal{E}_l = (\widetilde{I - A_1})^{-1} \widetilde{A_2} (X_{l-1} + \mathcal{E}_{l-1}) + (\widetilde{I - A_1})^{-1} \mathcal{F}_l.$$

Положим также

$$\begin{aligned} (\widetilde{I - A_1})^{-1} \widetilde{A_2} &= (I - A_1)^{-1} A_2 + \delta_1, \\ (\widetilde{I - A_1})^{-1} &= (I - A_1)^{-1} + \delta_2. \end{aligned}$$

Из (5.2.3) имеем $\mathbb{E} \delta_1 = \mathbb{E} \delta_2 = 0$. С учетом этих обозначений

$$X_l + \mathcal{E}_l = [(I - A_1)^{-1} A_2 + \delta_1] (X_{l-1} + \mathcal{E}_{l-1}) + (I - A_1)^{-1} \mathcal{F}_l + \delta_2 \mathcal{F}_l.$$

И поскольку X_l удовлетворяет соотношению (5.2.1), получим для $\mathcal{E}_l$ выражение

$$\mathcal{E}_l = (I - A_1)^{-1} A_2 \mathcal{E}_{l-1} + \delta_1 X_{l-1} + \delta_1 \mathcal{E}_{l-1} + \delta_2 \mathcal{F}_l \quad (5.2.4)$$

и

$$\mathcal{E}_l^T = \mathcal{E}_{l-1}^T [(I - A_1)^{-1} A_2]^T + \mathcal{E}_{l-1}^T \delta_1^T + X_{l-1}^T \delta_1^T + \mathcal{F}_l^T \delta_2^T. \quad (5.2.5)$$

Теперь мы должны перемножить правые и левые части равенств (5.2.4) и (5.2.5) соответственно и вычислить математические ожидания всех членов полученного равенства. При этом следует сразу же учесть, что математические ожидания многих слагаемых в правой части равны нулю.

Так, $E(\delta_1 X_{l-1} \mathcal{E}_{l-1}^T \delta_1^T) = 0$, так как $\mathcal{E}_{l-1}^T$ не зависит от δ_1 и $E \mathcal{E}_{l-1}^T = 0$, аналогично $E(\delta_1 \mathcal{E}_{l-1} \mathcal{F}_l^T \delta_2^T) = 0$. Учитывая эти соображения, имеем

$$E(\mathcal{E}_l \mathcal{E}_l^T) = (I - A_1)^{-1} A_2 E(\mathcal{E}_{l-1} \mathcal{E}_{l-1}^T) ((I - A_1)^{-1} A_2)^T + E(\delta_1 \mathcal{E}_{l-1} \mathcal{E}_{l-1}^T \delta_1^T) + \mathcal{F}_l,$$

где $\mathcal{F}_l$ — матрица, объединяющая слагаемые, не зависящие от $\mathcal{E}_{l-1}$.

Запишем далее

$$\text{cov} \mathcal{E}_l = (I - A_1)^{-1} A_2 \text{cov} \mathcal{E}_{l-1} ((I - A_1)^{-1} A_2)^T + E(\delta_1 \text{cov} \mathcal{E}_{l-1} \delta_1^T) + \mathcal{F}_l. \quad (5.2.6)$$

По определению операции умножения матриц найдем

$$\begin{aligned} \|C_{i,j}^{(l)}\|_{i,j-1}^n &= \left\| \sum_{i_1=1}^n \sum_{i_2=1}^n \alpha_{i_0,i_1} C_{i_1,i_2}^{(l-1)} \alpha_{i_3,i_2} \right\|_{i_0,i_3=1}^n + \\ &+ E \left\| \sum_{i_1=1}^n \sum_{i_2=1}^n \delta_{i_0,i_1}^l C_{i_1,i_2}^{(l-1)} \delta_{i_3,i_2}^l \right\|_{i_0,i_3=1}^n + \mathcal{F}_l, \end{aligned} \quad (5.2.7)$$

где $\alpha_{i,j}$ — элементы матрицы $(I - A_1)^{-1} A_2$, $C_{i,j}^{(l)}$ — матрицы $\text{cov} \mathcal{E}_l$ и $\delta_{i,j}^1$ — матрицы δ_1 . Введем мультииндекс $J = (i, j)$, принимающий n^2 значений от $(1, 1)$ до (n, n) . Тогда равенство (5.2.7) можно переписать в форме

$$\begin{aligned} \|C_J^{(l)}\| &= \left\| \sum_{J=(i_1,i_2)} \alpha_{i_0,i_1} \alpha_{i_3,i_2} C_J^{(l-1)} \right\| + \\ &+ \left\| \sum_{J=(i_1,i_2)} (E(\delta_{i_0,i_1} \delta_{i_3,i_2})) C_J^{(l-1)} \right\| + \mathcal{F}_l. \end{aligned} \quad (5.2.8)$$

Таким образом, если матрицу $\text{cov}\mathcal{E}_{l-1}$ “вытянуть” в один столбец (операция векторизации [33]) и трактовать далее $\|C_J^{(l)}\|$ как вектор с n^2 компонентами, то последнее равенство (5.2.8) можно записать в форме

$$\|C_J^{(l)}\| = \mathcal{L}\|C_J^{(l-1)}\| + \mathcal{M}\|C_J^{(l-1)}\| + \mathcal{F}_l,$$

где $\mathcal{L} = \|\alpha_{i_0, i_1} \alpha_{i_3, i_2}\|$ — матрица $n^2 \times n^2$, а $\mathcal{M} = \|\mathbb{E}(\delta_{i_0, i_1} \delta_{i_3, i_2})\|$ — матрица также $n^2 \times n^2$, i_0, i_1, i_2, i_3 принимают значения от 1 до n .

Справедливы следующие утверждения:

Утверждение 5.1. Собственные векторы $\mathcal{L}$ есть $\varphi_i \varphi_j^T$, где φ_i — собственный вектор (столбец) матрицы $(I - A_1)^{-1} A_2$; собственные числа $\mathcal{L}$ есть $\lambda_i \lambda_j$, где λ_i — собственные числа матрицы $(I - A_1)^{-1} A_2$, $i, j = 1, \dots, n$. $\square$

Утверждение 5.2. Элементы матрицы $\mathcal{M}$ есть ковариации вектора, составленного из элементов матрицы погрешностей δ_1 . Их величина имеет порядок $1/N$, где N — число независимых испытаний (траекторий) на $(l-1)$ -м шаге, то есть $\mathcal{M} = \frac{1}{N} \mathcal{M}'$, где $\mathcal{M}'$ — матрица $n^2 \times n^2$ с элементами, не зависящими от N . $\square$

Утверждение 5.1 проще всего проверить, учитывая, что $\|\alpha_{i,j}\| = (I - A_1)^{-1} A_2$; утверждение 5.2 вытекает из определения δ_1 .

Таким образом,

$$\|C_J^{(l)}\| = (\mathcal{L} + \frac{1}{N} \mathcal{M}') \|C_J^{(l-1)}\| + \mathcal{F}_l \quad (5.2.9)$$

и матрица ковариаций $\mathcal{E}_l$ будет оставаться ограниченной, если норма $\|\mathcal{L} + \frac{1}{N} \mathcal{M}'\|$ меньше единицы, и будет расти экспоненциально, если норма $\|\mathcal{L} + \frac{1}{N} \mathcal{M}'\|$ больше единицы. В последнем случае мы имеем дело со стохастической неустойчивостью. Случай равенства единице упомянутой нормы практически нереализуем.

Из (5.2.9) следуют утверждения.

Утверждение 5.3. Для стохастической устойчивости алгоритма (5.2.2) необходимо и достаточно, чтобы модуль первого собственного числа матрицы $(\mathcal{L} + \frac{1}{N} \mathcal{M}')$ был меньше единицы. $\square$

Утверждение 5.4. Если первое собственное число матрицы

$$(I - A_1)^{-1} A_2$$

строго меньше единицы, то существует натуральное число N_0 такое, что при всех $N > N_0$ алгоритм (5.2.2.) будет стохастически устойчив.

Это следует из утверждения 5.1 и непрерывной зависимости модуля первого собственного числа от величины элементов матрицы. $\square$

Вернемся теперь к разностному аналогу волнового уравнения. Рассматривается первая краевая задача на $[0, 1] \times [0, T]$. Явная схема в этом случае была стохастически неустойчивой. Рассмотрим простейшую неявную схему

$$\frac{u_{k,l+1} - 2u_{k,l} + u_{k,l-1}}{(\Delta t)^2} = \frac{u_{k+1,l+1} - 2u_{k,l+1} + u_{k-1,l+1}}{h^2}$$

или

$$u_{l+1,k+1} - (2 + \mu^2)u_{l+1,k} + u_{l+1,k-1} = \mu^{-2}(-2u_{l,k} + u_{l-1,k}). \quad (5.2.10)$$

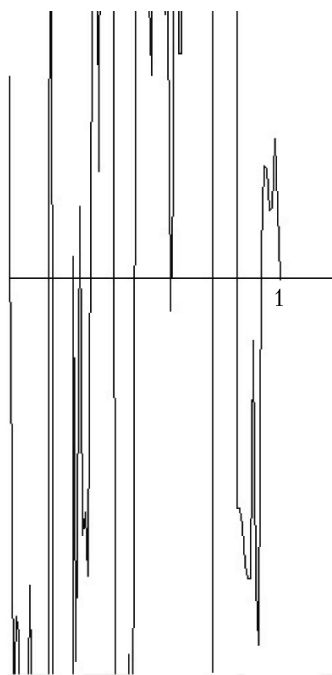
Как уже отмечалось (см. (5.1.7)), эта схема устойчива. Согласно утверждению 4, алгоритм вида (5.2.2) для системы уравнений (5.2.9) при некотором N_0 будет стохастически устойчив для всех $N \geq N_0$. Число N_0 зависит от m и μ^2 . Ниже приводятся графики, иллюстрирующие поведение решения уравнений (5.2.9) при некоторых значениях параметров $m, \Delta t$ и N .

На рис. 5.1 и 5.7 мы видим типичное неустойчивое поведение решения. Оно связано с явлением неустойчивости относительно ошибок округления при детерминированных вычислениях. На рис. 5.2 и 5.3 решение стохастически устойчиво, но явление неустойчивости может проявиться при больших L . На рис. 5.4–5.6 представлено стохастически устойчивое решение.

Отметим также, что в теоретическом плане случай зависимости искомой функции от нескольких пространственных переменных мало чем отличается от рассмотренного одномерного (по крайней мере для случая, когда область определения функции и есть прямоугольный гиперпараллелепипед). Если

$$L_h U_i = \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2}$$

приближает $\partial^2 U / \partial x^2$, то для многих переменных $\sum_{l=1}^s \partial^2 U / \partial x_l^2$ можно приблизить $\sum_{l=1}^s L_{h_l} U_{i_l}$. Легко проверить, что в рассматриваемом примере собственными векторами этого разностного оператора будут векторы, составленные из произведений соответствующих компонент собственных векторов операторов $L_{h_l} U_{i_l}$. Это позволяет провести полное исследование однородной первой краевой задачи. Результаты относительно устойчивости качественно сходны с одномерным случаем.



Число шагов по времени $L = 20$

Число точек деления

отрезка $(0, 1)$ $m = 100$

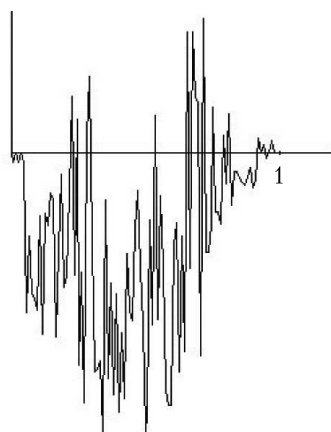
Шаг $h = 1/m$

Шаг по времени $\Delta t = 0,01$

Число независимых траекторий

на каждом слое $N = 2$

Рис. 5.1



Число шагов по времени $L = 10$

Число точек деления

отрезка $(0, 1)$ $m = 100$

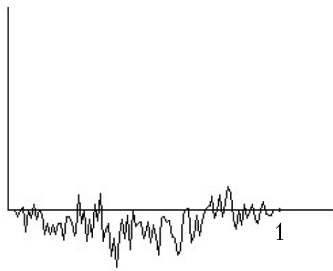
Шаг $h = 1/m$

Шаг по времени $\Delta t = 0,08$

Число независимых траекторий

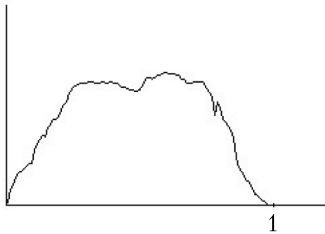
на каждом слое $N = 3$

Рис. 5.2



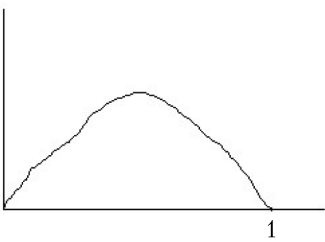
Число шагов по времени $L = 10$
Число точек деления
отрезка $(0, 1)$ $m = 100$
Шаг $h = 1/m$
Шаг по времени $\Delta t = 0,08$
Число независимых траекторий
на каждом слое $N = 5$

Рис. 5.3



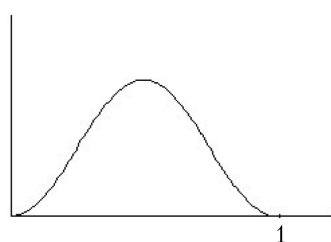
Число шагов по времени $L = 20$
Число точек деления
отрезка $(0, 1)$ $m = 100$
Шаг $h = 1/m$
Шаг по времени $\Delta t = 0,01$
Число независимых траекторий
на каждом слое $N = 10$

Рис. 5.4



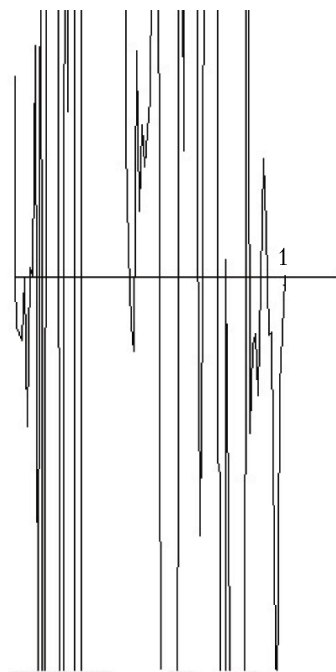
Число шагов по времени $L = 20$
Число точек деления
отрезка $(0, 1)$ $m = 100$
Шаг $h = 1/m$
Шаг по времени $\Delta t = 0,01$
Число независимых траекторий
на каждом слое $N = 50$

Рис. 5.5



Число шагов по времени $L = 3$
 Число точек деления
 отрезка $(0, 1)$ $m = 100$
 Шаг $h = 1/m$
 Шаг по времени $\Delta t = 0,01$
 Число независимых траекторий
 на каждом слое $N = 50$

Рис. 5.6



Число шагов по времени $L = 100$
 Число точек деления
 отрезка $(0, 1)$ $m = 100$
 Шаг $h = 1/m$
 Шаг по времени $\Delta t = 0,001$
 Число независимых траекторий
 на каждом слое $N = 5$

Рис. 5.7

5.3. Стохастическая устойчивость и параллелизм

На примере простейших разностных схем мы познакомились с явлением стохастической неустойчивости.

Для одномерного уравнения теплопроводности разностная схема

$$u_i^{t+1} = \frac{1}{2}(u_{i-1}^t + u_{i+1}^t)$$

стохастически устойчива. Случайное блуждание с началом в точке (T, i_0) , “спускаясь вниз”, укажет значение на границе, которое и будет несмещенной оценкой решения в этой точке. Здесь достаточно проследить траектории “по одной” и затем осреднить полученные результаты.

Для простейшего разностного аналога волнового уравнения

$$u_i^{t+1} = u_{i-1}^t + u_{i+1}^t - u_i^{t-1}$$

такой алгоритм построить нельзя. Эта разностная схема устойчива в обычном смысле (относительно ошибок округления), но неустойчива стохастически. Следует пользоваться неявной схемой, запоминая при каждом t вычисленные результаты. Здесь нужно моделировать пучок траекторий, запоминая результаты при каждом t . Если для каждого t делается N_1 испытаний и N_1 таково, что алгоритм стохастически устойчив (дисперсия не растет экспоненциально), то можно осуществить независимые вычисления N_2 раз (с N_1 повторениями) и результаты осреднить. Работает центральная предельная теорема, и можно построить доверительный интервал. Если N_1 таково, что алгоритм стохастически неустойчив, то этого делать нельзя: результаты быстро становятся бессмысленными и выходят за разрядную сетку.

Таким образом, при умеренных N_1 алгоритм метода Монте-Карло по-прежнему обладает свойством естественного параллелизма. Оценить N_1 можно в каждом случае в процессе вычислений.

Остановимся далее как на частных случаях, так и на обобщениях алгоритма, описанного в пп. 5.1 и 5.2.

5.3.1. Последовательные процедуры

Если в разбиении $A = A_1 + A_2$ мы полагаем $A_1 = 0$, то имеем итерационный процесс

$$X_{m+1} = AX_m + \mathcal{F}. \quad (5.3.1)$$

При этом из (5.1.16) следует, что на $A = A_2$ налагается условие $|\lambda_1(A)| < 1$ (но не $\lambda_1(|A|) < 1$!). Здесь схема Неймана–Улама может оказаться неприменимой.

В случае если $\lambda_1(|A|) > 1$, но $|\lambda_1(A)| < 1$, можно воспользоваться схемой с запоминанием промежуточных результатов. Эта процедура бывает весьма эффективной, если порядок системы n очень велик. Для наших целей необходимо указать алгоритм построения последовательности независимых случайных векторов $\Xi_j, j = 0, \dots, m$, таких что выполняется равенство

$$E(\Xi_m | \Xi_{m-1}) = A\Xi_{m-1} + \mathcal{F}. \quad (5.3.2)$$

Сделать это можно многими способами. Ниже описывается один из таких алгоритмов.

Полагаем $\Xi_0 = \mathcal{F}$. Выбираем распределение p^0 , согласованное с $\mathcal{F}$, и стохастическую матрицу $\mathcal{P}$, согласованную с A^T . Моделируем p_0 , находим i_{m-1} . Моделируем распределение $p_{i_{m-1},1}, \dots, p_{i_{m-1},n}$ (i_{m-1} -я строка матрицы $\mathcal{P}$), находим i_m . Строим случайный вектор η_m , все компоненты которого, кроме компоненты с номером i_m , равны нулю. i_m -я компонента полагается равной

$$\frac{\xi_{i_{m-1}}^{m-1}}{p_{i_{m-1}}^0} \cdot \frac{a_{i_m, i_{m-1}}}{p_{i_{m-1}, i_m}}, \quad \text{где } \Xi_{m-1} = (\xi_1^{m-1}, \dots, \xi_n^{m-1}) \quad \text{и} \quad \mathcal{P} = \|p_{i,j}\|.$$

Вычисляем N_1 реализаций η_m^j вектора η_m . Полагаем

$$\Xi_m = \frac{1}{N_1} \sum_{j=1}^{N_1} \eta_m^j + \mathcal{F}. \quad (5.3.3)$$

Легко проверить, что для определенного таким образом Ξ_m выполняется равенство (5.3.2).

К последовательностям Ξ_m , для которых выполняется равенство (5.3.2), применима теория предыдущих разделов гл. 5. Для рассмотренного случая справедлив общий вывод, что при $|\lambda_1(A)| \leq 1 - \delta$, $\delta > 0$ найдется N_1 такое, при котором алгоритм будет стохастически устойчив.

5.3.2. Методы *квази*Монте-Карло

Векторам X и $\mathcal{F}$ всегда могут быть сопоставлены функции, имеющие проекции $x_1, \dots, x_n$ и $f_1, \dots, f_n$ соответственно на ортонормированные базисные функции. Матрицу A также можно рассматривать как

оператор из n -мерного пространства функций в n -мерное. В частности, в качестве базисных могут быть выбраны ступенчатые функции на $[0, 1]$.

Это тривиальное замечание позволяет после простых рассуждений сделать вывод, что оценки решений систем линейных алгебраических уравнений (см. гл. 4) можно считать функциями ограниченной вариации на пространстве траекторий и применять квазислучайные числа для решения упомянутых систем (см. гл. 3).

В квазислучайной схеме траектория цепи Маркова длины d потребует вычисления координат d -мерной квазислучайной точки. При больших d процедура может быть излишне трудоемкой. Один из способов обойти эту трудность состоит в том, чтобы использовать квазислучайные координаты только до некоторой фиксированной длины $d \leq s$, а далее прибегнуть к датчику псевдослучайных чисел. Константная размерность вычисляемого интеграла равна s , и оценка погрешности дается выражением

$$C_1 \frac{(\ln N)^s}{N} + C_2 \frac{1}{\sqrt{N}}, \quad (5.3.4)$$

где C_2 — малая величина при удачно подобранном s . Величина C_2 тем меньше, чем меньше количество траекторий, длина которых превышает s . Выражение (5.3.4) при больших N дает лучшую величину погрешности, чем обычный метод Монте-Карло. Однако при больших s это N может быть слишком большим.

Другой подход основан на использовании метода, описанного в схеме *квази*Монте-Карло (см. п. 5.3.1).

Поскольку при каждом m оценивается интеграл кратности 1, погрешность убывает как N^{-1} . Нужно только, чтобы при этом N обеспечивалась стохастическая устойчивость алгоритма. Требуется согласование с величиной δ в неравенстве $|\lambda_1(A)| \leq 1 - \delta$. Вопрос о выборе метода решается с помощью предварительных численных экспериментов.

5.4. Обобщение на случай операторов

Далее обобщение некоторых из полученных результатов распространяется на рекуррентную процедуру^{*)}

$$\varphi^m = A_0 \varphi^m + A_1 \varphi^{m-1} + f^m, \quad m \geq 1, \quad (5.4.1)$$

^{*)} В отличие от предыдущих разделов теперь m будет обозначать номер итерации.

где A_0, A_1 — линейные интегральные операторы; φ^m, f^m — функции, заданные на множестве $D \subset \mathbf{R}^s$. Предложен рекуррентный алгоритм для получения последовательностей оценок $\hat{\varphi}^m$ функций φ^m , использующий для аппроксимации функций на каждом шаге процедуру непараметрического оценивания. Получены достаточные условия стохастической устойчивости рассматриваемого алгоритма [5]. Заметим, что поскольку оценки $\hat{\varphi}^m$ являются смещенными, то для устойчивости алгоритма кроме условия

$$\|\text{cov}(\hat{\varphi}^m)\| < C \quad \text{для всех } m \geq 1$$

необходимо, чтобы систематическая ошибка не увеличивалась с ростом m .

5.4.1. Схема Неймана–Улама

Предварительно напомним некоторые сведения, касающиеся схемы Неймана–Улама (см. гл. 4). Обозначим

$$A\varphi(x) = \int a(x, y)\varphi(y)\mu(dy), \quad x, y \in D \subset \mathbf{R}^d,$$

где μ — σ -конечная мера, $\varphi \in \mathbf{F}$, $\mathbf{F} \subset \mathbf{L}^2(d\mu)$, и $\langle f, g \rangle$ означает скалярное произведение функций $f, g \in \mathbf{F}$.

Как известно, схема Неймана–Улама применима в случае

$$\|\bar{A}\| < 1, \quad \text{где } \bar{A}\varphi(x) = \int |a(x, y)|\varphi(y)\mu(dy). \quad (5.4.2)$$

Если условие (5.4.2) выполнено, то можно выбрать однородную цепь Маркова и функционал (оценку) $\xi = \xi(\varphi, h)$ на ее траекториях такие, что

$$\mathbf{E} \xi(\varphi, h) = \langle \varphi, h \rangle \quad \text{и} \quad \mathbf{E} \xi^2(\varphi, h) < +\infty. \quad (5.4.3)$$

Процедура аппроксимации функции φ в D , которую мы будем рассматривать далее, состоит в следующем. Реализуются алгоритмы метода Монте-Карло в схеме Неймана–Улама для оценки значений величин $\{\langle \varphi, h_i \rangle, i = 1, \dots, N\}$, а затем по полученным приближенным значениям $\{\xi(\varphi, h_i)\}$ с помощью некоторого преобразования G происходит восполнение функции φ . В итоге оценка функции φ выглядит следующим образом:

$$\hat{\varphi}(x) = G(\xi(\varphi, h_1), \dots, \xi(\varphi, h_N), x), \quad x \in D. \quad (5.4.4)$$

Частным случаем оценки (5.4.4) является оценка ядерного типа

$$\widehat{\varphi}(x) = \frac{1}{N} \sum_{i=1}^N K(x, x_i) y(x_i) / q(x_i), \quad (5.4.5)$$

где $K(x, z)$ — функция, заданная на множестве $D \times D$ и близкая к δ -функции $\delta(x - z)$; $x_1, \dots, x_N$ — независимые реализации случайного вектора x в D , имеющего плотность распределения $q(x)$; $y(x_i)$ — оценка (по поглощению) значения функции φ в точке x_i .

Поскольку φ является решением уравнения $\varphi = A\varphi + f$ (где A — линейный интегральный оператор; функции φ, f — элементы некоторого банахова пространства $\mathbf{F}$, $h \in \mathbf{F}^*$), то случайный оператор — обозначим его $(\widetilde{I - A_0})^{-1}$, — сопоставляющий функции f оценку $\widehat{\varphi}$, можно рассматривать в качестве оценки оператора $(I - A_0)^{-1}$.

Напомним, что существует важный класс задач, для которого схема Неймана–Улама также неприменима на практике, несмотря на то что все условия ее использования формально выполнены. Действительно, рассмотрим итерационный процесс

$$\varphi^m = A_0 \varphi^m + A_1 \varphi^{m-1} + f^m, \quad m \geq 1, \quad (5.4.6)$$

где A_0, A_1 — линейные интегральные операторы.

Введем матрицу-оператор размерности $m \times m$

$$\mathcal{A}^{(m)} = \begin{pmatrix} A_0 & A_1 & 0 & \dots & 0 & 0 \\ 0 & A_0 & A_1 & \dots & 0 & 0 \\ 0 & 0 & A_0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & A_0 & A_1 \\ 0 & 0 & 0 & \dots & 0 & A_0 \end{pmatrix}$$

и вектор-функции размерности m

$$F^{(m)} = \begin{pmatrix} f^m \\ f^{m-1} \\ \vdots \\ f^2 \\ A_0 \varphi^0 + f^1 \end{pmatrix}, \quad \phi^{(m)} = \begin{pmatrix} \varphi^m \\ \varphi^{m-1} \\ \vdots \\ \varphi^2 \\ \varphi^1 \end{pmatrix}, \quad H^{(m)} = \begin{pmatrix} h \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

Тогда итерационный процесс (5.4.6) записывается в виде

$$\phi^{(m)} = \mathcal{A}^{(m)} \phi^{(m)} + F^{(m)}, \quad m \geq 1. \quad (5.4.7)$$

Применяя схему Неймана–Улама к уравнению (5.4.7), можно получить оценки

$$\xi = \xi(\phi^{(m)}, H^{(m)}) \text{ такие, что } E\xi = \langle \varphi^m, h \rangle \text{ и } E\xi^2 < +\infty,$$

и затем построить оценку $\hat{\varphi}^m$ функции φ^m .

Заметим, что условие применимости схемы Неймана–Улама (5.4.2) для уравнения (5.4.7) сводится к условию $\|A_0\| < 1$, и мы формально можем прибегнуть к ней, даже если $\|A_1\| > 1$. Однако в последнем случае возможна ситуация, когда ковариация оценки $\hat{\varphi}^m$, будучи конечной при фиксированном m , с ростом m будет расти экспоненциально. Поэтому описанный выше метод становится неприемлемым с ростом m .

Следовательно, если мы хотим строить оценки $\hat{\varphi}^m$ для больших m , необходимо, чтобы метод был стохастически устойчивым, то есть выполнялось условие

$$\|\text{cov}(\hat{\varphi}^m)\| < +\infty \text{ при } m \rightarrow \infty.$$

5.4.2. Рекуррентный алгоритм

Рассмотрим рекуррентную процедуру

$$\varphi^m = A_0\varphi^{m-1} + A_1\varphi^{m-2} + f^m, \quad m \geq 1, \quad (5.4.8)$$

где $\varphi^0, f^m \in \mathbf{F}$ — заданные функции переменной x , $x \in D$.

Будем строить рекуррентную последовательность оценок $\hat{\varphi}^m$ для функций φ^m , удовлетворяющих соотношению (5.4.8).

Положим

$$\hat{\varphi}^0(x) = \varphi^0(x).$$

Пусть получена оценка $\hat{\varphi}^{m-1}$ функции φ^{m-1} . На m -м шаге строим операторные оценки $((\widetilde{I - A_0})^{-1} A_1)^m$ и $((\widetilde{I - A_0})^{-1})^m$ с помощью процедуры, описанной в п. 5.4.1, и полагаем

$$\hat{\varphi}^m(x) = ((\widetilde{I - A_0})^{-1} A_1)^m \hat{\varphi}^{m-1}(x) + ((\widetilde{I - A_0})^{-1})^m f^m(x). \quad (5.4.9)$$

Теорема 5.1. (Обобщение (5.2.9).) Пусть выполняется условие

$$\|(I - A_0)^{-1} A_1\| < 1. \quad (5.4.10)$$

Тогда существует число N_0 такое, что при любом $N > N_0$ рекуррентная процедура (5.4.9) стохастически устойчива.

Доказательство. Введем обозначение для ошибки:

$$\varepsilon^m(x) = \widehat{\varphi}^m(x) - \varphi^m(x)$$

и исследуем поведение первого и второго моментов ошибки ε^m с ростом m . Имеем

$$\mathbb{E}\varepsilon^m(x) = \mathbb{E}\widehat{\varphi}^m(x) - \varphi^m(x) = \delta_2^m(x), \quad (5.4.11)$$

где δ_2^m — систематическая погрешность на m -м шаге.

Через $\mathcal{F}_m$ обозначим σ -алгебру, порожденную $\widehat{\varphi}^m$; через $\mathcal{F}_h$ — σ -алгебру, порожденную h_i , $1 \leq i \leq N$; $\mathbb{E}_m$ и $\mathbb{E}_h$ — условные математические ожидания относительно $\mathcal{F}_m$ и $\mathcal{F}_h$ соответственно; δ_1 — оператор ошибки восполнения:

$$\delta_1\varphi(x) = G(\langle\varphi, h_1\rangle, \dots, \langle\varphi, h_N\rangle, x) - \varphi(x).$$

Используя тот факт, что G является линейным преобразованием, имеем

$$\begin{aligned} \mathbb{E}\widehat{\varphi}^m(x) &= \mathbb{E}G(\xi(\varphi, h_i), 1 \leq i \leq N, x) = \\ &= \mathbb{E}G(\mathbb{E}_h\mathbb{E}_{m-1}\xi(\varphi, h_i), 1 \leq i \leq N, x) = \\ &= \mathbb{E}G(\mathbb{E}_h\langle(I - A_0)^{-1}A_1\widehat{\varphi}^{m-1} + (I - A_0)^{-1}f^m, h_i\rangle, 1 \leq i \leq N, x) = \\ &= \mathbb{E}G(\langle(I - A_0)^{-1}A_1(\varphi^{m-1} + \delta_2^{m-1}) + (I - A_0)^{-1}f^m, h_i\rangle, 1 \leq i \leq N, x) = \\ &= \varphi^m(x) + \delta_1\varphi^m(x) + (I - A_0)^{-1}A_1\delta_2^{m-1}(x) + \delta_1(I - A_0)^{-1}A_1\delta_2^{m-1}(x). \end{aligned}$$

Отсюда и из равенства (5.4.11) получаем рекуррентное соотношение для погрешностей

$$\delta_2^m(x) = ((I - A_0)^{-1}A_1 + \delta_1(I - A_0)^{-1}A_1)\delta_2^{m-1}(x) + \delta_1\varphi^m(x). \quad (5.4.12)$$

□

Замечание. В случае оценки ядерного типа (5.4.5) оператор δ_1 имеет вид

$$\delta_1\varphi(x) = \frac{1}{N} \sum_{i=1}^N \mathbb{E}K(x, x_i)\varphi(x_i)/q(x_i) - \varphi(x) = \int \varphi(z)K(x, z)dz - \varphi(x). \quad (5.4.13)$$

□

Из (5.4.13) следует, что если функция $K(x, z)$ близка к δ -функции $\delta(x - z)$ и функция φ удовлетворяет условию Липшица на множестве D , то величина $\delta_1 \varphi(x)$ равномерно близка к 0 при всех $x \in D$. Относительно выполнения будем предполагать, что $\|\delta_1\|$ стремится к 0 при $N \rightarrow \infty$. Тогда, учитывая условие (5.4.10), можно утверждать, что найдется такое число N_1 , что выполняется неравенство

$$(1 + \|\delta_1\|) \|(I - A_0)^{-1} A_1\| < 1. \quad (5.4.14)$$

Из соотношения (5.4.12) имеем

$$\sup_m \|\delta_2^m\| \leq C \sup_m \|\varphi^m\|, \quad \text{где } C = \frac{\|\delta_1\|}{1 - (1 + \|\delta_1\|) \|(I - A_0)^{-1} A_1\|}.$$

Таким образом, малость нормы ошибки выполнения обеспечивает малость нормы систематической ошибки.

Считая ошибку выполнения малой, будем анализировать поведение вторых моментов случайной ошибки. Обозначим

$$R^m(x, y) = \text{cov}(\varepsilon^m(x), \varepsilon^m(y)).$$

Покажем, что величина R^m подчиняется некоторому рекуррентному соотношению

$$R^m(x, y) = \mathcal{U}R^{m-1}(x, y) + \rho^m(x, y), \quad (5.4.15)$$

причем для оператора $\mathcal{U} = \mathcal{U}_N$ выполняется условие $\|\mathcal{U}\| < 1$ при $N > N_0$.

Очевидно, что

$$R^m(x, y) = \text{cov}(\varepsilon_1^m(x), \varepsilon_1^m(y)),$$

где

$$\varepsilon_1^m(x) = \widehat{\varphi}^m(x) - \mathbf{E} \widehat{\varphi}^m(x).$$

Для случайной ошибки ε_1^m выполняется

$$\begin{aligned} \varepsilon_1^m(x) &= ((I - \widetilde{A_0})^{-1} A_1)^m \varepsilon_1^{m-1}(x) + ((I - \widetilde{A_0})^{-1} A_1)^m \mathbf{E} \widehat{\varphi}^{m-1}(x) + \\ &+ ((I - \widetilde{A_0})^{-1})^m f^m(x) - \mathbf{E} \widehat{\varphi}^m(x) = \Delta_1^m \mathbf{E} \widehat{\varphi}^{m-1}(x) + \Delta_2^m f^m(x) + \\ &+ (I - A_0)^{-1} A_1 \varepsilon_1^{m-1}(x) + \delta_1 (I - A_0)^{-1} A_1 \varepsilon_1^{m-1}(x) + \Delta_1^m \varepsilon_1^{m-1}(x), \end{aligned} \quad (5.4.16)$$

где случайные операторы Δ_1^m и Δ_2^m имеют вид

$$\begin{aligned}\Delta_1^m &= ((I - \widetilde{A_0})^{-1}A_1)^m - E_{m-1}((I - \widetilde{A_0})^{-1}A_1)^m, \\ \Delta_2^m &= ((I - \widetilde{A_0})^{-1})^m - E_{m-1}((I - \widetilde{A_0})^{-1})^m.\end{aligned}$$

Учитывая, что $E_{m-1}\Delta_i^m = 0$, $i = 1, 2$, и $E\varepsilon_1^m(x) = 0$, из (5.4.16) получаем *)

$$\begin{aligned}R^m(x, y) &= \text{cov}((I + \delta_1)(I - A_0)^{-1}A_1\varepsilon_1^{m-1})(x, y) + \\ &+ \text{cov}(\Delta_1^m\varepsilon_1^{m-1})(x, y) + \rho^m(x, y),\end{aligned}\quad (5.4.17)$$

где

$$\rho^m(x, y) = \text{cov}(\Delta_1^m E\widehat{\varphi}^{m-1} + \Delta_2^m f^m)(x, y).$$

Затем для интегрального оператора A определим оператор $\mathcal{M}(A)$:

$$\mathcal{M}(A)r(x, y) = \iint a(x, u)r(u, v)a(y, v)\mu(du)\mu(dv),$$

где функция $r \in \mathbf{L}^2(d\mu \otimes d\mu)$, тогда

$$\text{cov}((I + \delta_1)(I - A_0)^{-1}A_1\varepsilon_1^m)(x, y) = \mathcal{M}((I + \delta_1)(I - A_0)^{-1}A_1)R^{m-1}(x, y). \quad (5.4.18)$$

Замечание. Используя неравенство Гёльдера, нетрудно показать, что имеет место равенство $\|\mathcal{M}(A)\| = \|A\|^2$. $\square$

Далее будем рассматривать такие восполнения, для которых выполняются следующие условия: существует линейный оператор $\mathcal{L} = \mathcal{L}_N$ такой, что для любой функции $g \in \mathbf{F}$ имеет место равенство

$$E_{m-1}\Delta_1^m g(x)\Delta_1^m g(y) = \mathcal{L}g(x)g(y) \quad (5.4.19)$$

и

$$\|\mathcal{L}\| \rightarrow 0 \text{ при } N \rightarrow \infty. \quad (5.4.20)$$

Замечание. Можно показать, что оценка ядерного типа (5.4.5) обладает свойствами (5.4.19)–(5.4.20). $\square$

Из (5.4.17)–(5.4.19) получаем

$$R^m(x, y) = (\mathcal{M}((I + \delta_1)(I - A_0)^{-1}A_1) + \mathcal{L})R^{m-1}(x, y) + \rho^m(x, y).$$

*) Здесь и далее для случайной функции $\varepsilon(x)$ используется сокращенное обозначение $\text{cov}(\varepsilon)(x, y) \equiv \text{cov}(\varepsilon(x), \varepsilon(y))$.

Таким образом, оператор $\mathcal{U}$ из рекуррентного соотношения (5.4.15) имеет вид

$$\mathcal{U} = \mathcal{M}((I + \delta_1)(I - A_0)^{-1}A_1) + \mathcal{L}.$$

Принимая во внимание условия (5.4.14) и (5.4.20), а также последнее замечание, имеем

$$\|\mathcal{M}((I + \delta_1)(I - A_0)^{-1}A_1)\| = \|(I + \delta_1)(I - A_0)^{-1}A_1\|^2 < 1,$$

и из условия $\|(I - A_0)^{-1}A_1\| < 1$ при достаточно большом N следует $\|\mathcal{U}\| < 1$.

5.5. Библиографическое дополнение к главе 5

Вопросы восполнения (интерполяции) оцениваемой функции подробно изучались, в частности, Войтишеком и Михайловым [59]. В нашем изложении используются известные факты из теории непараметрического оценивания функции регрессии. Точки, в которых оценивается искомая функция, не предполагаются фиксированными. Это может оказаться важным при решении ряда задач. Во всех случаях погрешность имеет порядок $N^{1/2}$, улучшение возможно за счет использования метода зависимых испытаний [80]. Систематическое исследование этого метода в теории ядерных оценок не проводилось.

Стохастическая устойчивость в связи с разностными схемами изучалась в работах [32], [37], [38]. Достаточные условия, близкие к необходимым, впервые получены Вагнером и Ермаковым [7]. В работе [30] в связи с этим обсуждаются проблемы параллелизма алгоритмов. Калошиным составлена программа, позволяющая проследить нарушение стохастической устойчивости при решении разностного аналога волнового уравнения [51]. Рисунки 5.1–5.7 получены с помощью алгоритма, рассматриваемого в [17].

Алгоритм последовательного оценивания итераций, описанный в п. 5.3.1, был использован Ермаковым и Рукавишниковой [47]. Близкий к п. 5.3.1 алгоритм был также независимо использован Сабельфельдом при решении краевых задач [65].

Глава 6

Уравнения с полиномиальной нелинейностью

Ранее (см. гл. 4 и 5) нами были рассмотрены некоторые стохастические методы решения линейных уравнений, основанные на сопоставлении итерационных методов и свойств цепей Маркова.

Нелинейные задачи обладают большим разнообразием. В вычислительной математике решение нелинейных задач, как правило, сводится тем или иным методом к решению линейных уравнений. Широко известны метод Ньютона–Канторовича и ряд других методов последовательных приближений, требующих на каждом этапе решения линейной задачи.

В таком подходе для решения линейных задач может быть использован метод Монте-Карло, и, подобно задачам предыдущей главы, мы будем иметь систематическую и случайную погрешности, поведение которых необходимо исследовать. Оказывается также, что для некоторых классов уравнений удастся построить алгоритм, основанный на моделировании марковских цепей, и указать класс оценок функционалов от решений, используя схему Неймана–Улама.

Прежде всего мы рассмотрим уравнения именно такого класса — уравнения с полиномиальной нелинейностью. С ними связаны ветвящиеся марковские процессы, и оценки их решений можно представить в виде функционалов от ветвящихся траекторий (деревьев). Известно, что производящая функция ветвящегося процесса — многочлен (оператор с полиномиальной нелинейностью в общем случае) [67]. С другой стороны, так называемые суммы Вальда являются аналогом ряда Неймана для полиномиальных операторов. В п. 6.1 схематично намечена связь ветвящихся процессов и уравнений с полиномиальной нелинейностью. Это позволяет описать алгоритм моделирования, построить аналог оценки по поглощениям $\mathfrak{J}_a[\omega]$ и исследовать ее дисперсию.

Далее, в п. 6.2, используется известная связь ветвящихся процессов и процесса блуждания на прямой. Формально эта связь может быть выражена путем последовательного домножения обеих частей уравнения на искомую функцию от новой независимой переменной [61]. Применение схемы Неймана–Улама к полученной бесконечной линейной системе позволяет построить в нелинейном случае аналоги известных для линейного случая несмещенных оценок [6].

6.1. Ветвящиеся процессы и нелинейные уравнения

Поясним сначала основную идею этого раздела на примере простейшей квадратичной нелинейности. Изложение в случае произвольной полиномиальной нелинейности неизбежно громоздко, и доказательство леммы 6.1 может быть опущено читателем, который не заинтересовался специально этим узким вопросом. Отметим, что подход, предложенный в п. 6.2, более прозрачен.

По-видимому, простейшей нелинейностью, заслуживающей исследования, является квадратичная. Рассмотрим уравнение

$$\varphi(x) = \int \int k(x, y, z) \varphi(y) \varphi(z) \mu(dx) \mu(dy) + f(x) \pmod{\mu}, \quad (6.1.1)$$

или, в сокращенной записи,

$$\varphi = \mathcal{K}\varphi\varphi + f \quad (6.1.2)$$

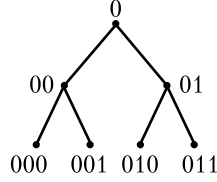
(существование соответствующих интегралов предполагается). Положим $\varphi_0 = f$ и определим функцию φ_n рекуррентно:

$$\varphi_{n+1} = \mathcal{K}\varphi_n\varphi_n + f. \quad (6.1.3)$$

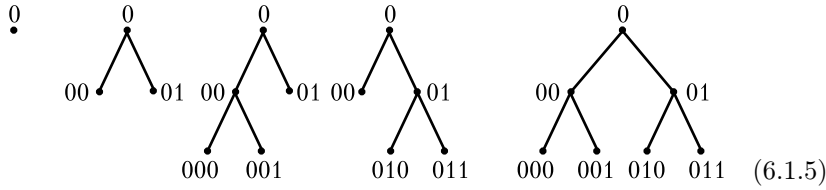
Имеем также

$$\begin{aligned} \varphi_{n+1} &= f + \mathcal{K}(f + \mathcal{K}\varphi_{n-1}\varphi_{n-1})(f + \mathcal{K}\varphi_{n-1}\varphi_{n-1}) = \\ &= f + \mathcal{K}ff + \mathcal{K}f(\mathcal{K}\varphi_{n-1}\varphi_{n-1}) + \mathcal{K}(\mathcal{K}\varphi_{n-1}\varphi_{n-1})f + \\ &\quad + \mathcal{K}(\mathcal{K}\varphi_{n-1}\varphi_{n-1})(\mathcal{K}\varphi_{n-1}\varphi_{n-1}). \end{aligned} \quad (6.1.4)$$

Рассмотрим теперь двоичное дерево вида



и множество его всевозможных поддеревьев, то есть



Дерево и его поддеревья состоят из “троек” (на языке частиц — из частицы и двух ее потомков, например: 01, 010, 011). Исключение составляет начальная точка (вырожденное дерево).

Как видим, между слагаемыми в (6.1.4) и множеством поддеревьев (6.1.5) существует взаимно однозначное соответствие. Более того, каждому $\mathcal{K}$ соответствует “тройка”, частицам с трехзначными номерами — φ_{n-1} , а начальной точке и точкам с двузначными номерами — f . (При $n = 1$ функция f соответствует всем частицам, не имеющим потомков.)

Для того чтобы разобраться, как φ_{n+1} выражается через f , следует ввести понятие полного двоичного дерева с $n + 1$ поколением. Формально его можно определить как упорядоченную совокупность целых двоичных чисел, имеющих не менее чем $n + 1$ двоичную цифру, причем первой цифрой является нуль. Поддерево такого дерева мы определим как подмножество чисел, входящих в полное дерево и обладающих следующим свойством.

Если число $r0 = 0\varepsilon_1 \dots \varepsilon_k 0$ принадлежит поддереву γ_n , $k < n - 1$, то ему принадлежат также числа $r1 = 0\varepsilon_1 \dots \varepsilon_k 1$ и $r = 0\varepsilon_1 \dots \varepsilon_k$ ($\varepsilon_j = 0, 1; j = 1, \dots, k$). Множество всех поддеревьев полного дерева с n поколениями будем обозначать Γ_n .

Каждому поддереву γ_n сопоставим функцию переменных $x(r)$ (x с номером r), где r — номера, входящие в поддерево γ_n , которое строится следующим образом. Каждой тройке $(r, r0, r1)$ сопоставляем

$k(x(r), x(r0), x(r1))I(x(r0))I(x(r1))$, где $k(x, y, z)$ — ядро оператора $\mathcal{K}$, а $I(x)$ — функция, определяемая равенством (напомним, что $\varphi_0 = f$):

$$I(x) = \begin{cases} f(x), & \text{если } x \text{ — концевая точка дерева} \\ & \text{и } x \text{ принадлежит поколению с номером,} \\ & \text{меньшим } n, \\ 1, & \text{если } x \text{ — координата частицы,} \\ & \text{имеющей потомков,} \\ \varphi_s(x) \ (s \geq 0), & \text{если частица принадлежит поколению} \\ & \text{с номером } n; \end{cases} \quad (6.1.6)$$

Затем на γ_n определим функцию $\mathfrak{J}(\gamma_n)$:

$$\mathfrak{J}(\gamma_n) = \prod_{\gamma_n} k(x(r), x(r0), x(r1))I(x(r0))I(x(r1)), \quad (6.1.7)$$

где символ $\prod_{\gamma_n}$ означает произведение по всем тройкам $(r, r0, r1)$ из γ_n .

Лемма 6.1. Справедливо равенство

$$\varphi_{n+s}(x) = \sum_{\gamma_n \in \Gamma_n} \int \mathfrak{J}(\gamma_n) d\mu^{|\gamma_n|}, \quad (6.1.8)$$

где суммирование ведется по всем поддеревьям из Γ_n , а $d\mu^{|\gamma_n|}$ означает интегрирование по всем переменным с номерами из γ_n , кроме x . **Доказательство.** Доказательство проведем методом индукции по n . Базу индукции задают равенствами (6.1.3), (6.1.4). Предположим, что равенство (6.1.7) имеет место для некоторого фиксированного $n = n_0$. Покажем, что оно выполняется для $n_0 + 1$. Справедливо равенство

$$\varphi_{n_0+s+1} = \sum_{\gamma_{n_0+1} \in \Gamma_{n_0+1}} \int \mathfrak{J}(\gamma_{n_0+1}) d\mu^{|\gamma_{n_0+1}|}. \quad (6.1.9)$$

Действительно, пусть $r0$ и $r1$ — некоторые номера координат частиц n -го поколения и $s > 0$; $r0, r1$ являются $(n+1)$ -значными двоичными числами. Выразим $\varphi_1(x)$ через $\varphi_{s-1}(x)$ в соответствии с (6.1.3), нумеруя переменные специальным образом. Имеем

$$\begin{aligned} \varphi_s(xr0) &= f(xr0) + \int \int k(xr0, xr00, xr01) \varphi_{s-1}(xr00) \times \\ &\quad \times \varphi_{s-1}(xr01) \mu(dxr00) \mu(dxr01) \varphi_s(xr1) = \\ &= f(xr1) + \int \int k(xr1, xr10, xr11) \varphi_{s-1}(xr10) \varphi_{s-1}(xr11) \mu(dxr10) \mu(dxr11). \end{aligned} \quad (6.1.10)$$

Заметим теперь, что каждое γ_{n_0} может предшествовать некоторому γ_{n_0+1} , которое образуется следующим образом: у каждой частицы n -го поколения из γ_n с номерами $r0$ и $r1$ могут появиться потомки — $r00, r01, r10, r11$ соответственно. Каждой комбинации этих (четырёх) возможностей для каждой из комбинаций частиц n -го поколения соответствует дерево γ_{n_0+1} из Γ_{n_0+1} .

Перемножая $\varphi_1(xr0)$ и $\varphi_1(xr1)$ (правые части равенств (6.1.10)), мы видим, что имеется ровно одно слагаемое, соответствующее каждой из перечисленных комбинаций. Это и доказывает лемму. $\square$

При $s = 1$, $\varphi_{s-1} = \varphi_0 = f$ можно определить $I(x)$ в (6.1.6) следующим образом:

$$I(x) = \begin{cases} f(x), & \text{если } x \text{ — координата частицы, не имеющей потомков,} \\ 1, & \text{если } x \text{ — координата частицы, имеющей потомков;} \end{cases}$$

здесь считается, что частицы $(n+1)$ -го поколения не имеют потомков.

Сформулируем следствие леммы 6.1.

Следствие. $(n+1)$ -я итерация φ_{n+1} выражается равенством

$$\varphi_{n+1}(x_0) = \sum_{\gamma_{n+1} \in \Gamma_n} \int \mathfrak{J}(\gamma_n d\mu^{|\gamma_n|}), \quad (6.1.11)$$

где $\mathfrak{J}(\gamma_{n+1})$ определяется равенством (6.1.7). $\square$

Исследование итераций уравнения (6.1.1) с простейшей нелинейностью потребовало введения довольно сложных обозначений. По этой причине мы без доказательств укажем на связи деревьев более общей структуры с уравнениями вида

$$\varphi(x) = \sum_{l=1}^M \int d\mu^l k_l(x, x_1, \dots, x_l) \prod_{j=1}^l \varphi(x_j) + f(x) \pmod{\mu}.$$

Имеет место результат, аналогичный полученному нами ранее для уравнения (6.1.1). Итерации оператора в правой части (6.1.11) также связаны с поддеревьями, в которых каждая частица может иметь $0, 1, \dots, l$ потомков. Точнее, каждая частица обладает дополнительной характеристикой — “сортом”. Частица s -го сорта ($s = 0, 1, \dots, M$) может иметь s потомков. Каждый потомок несет информацию о сорте предка и о своем номере среди s потомков. $\mathfrak{J}(\gamma_n)$ в этом случае имеет

вид

$$\mathfrak{J}(\gamma_n) = \prod_{\gamma_n} k_s(x_0, x^{s(r)}(r, 0), \dots, x^{s(r)}(r, s-1)) \times \prod_{j \in \Delta(\gamma_n)} f(x(j)), \quad (6.1.12)$$

где $\Delta(\gamma_n)$ — множество номеров точек из γ_n , не имевших потомков.

Как и ранее, справедливо равенство

$$\varphi_n(x) = \sum_{\gamma_n \in \Gamma_n} \int \mathfrak{J}(\gamma_n) d\mu^{|\gamma_n|}. \quad (6.1.13)$$

Ветвящиеся цепи Маркова

Рассмотрим теперь случайные процессы, порождающие деревья (в принятой ранее терминологии — поддеревья) описанного выше типа. Пусть время t принимает дискретные значения: $t = 0, 1, \dots$. Однородный процесс рождения/гибели частиц зададим плотностью (по отношению к мере μ) начального распределения $p^0(x)$ и плотностями (по отношению к μ^s) перехода с рождением s частиц ($s \geq 1$):

$$p_s(x, x_1, \dots, x_s).$$

При этом должно выполняться условие нормировки

$$\sum_{s=0}^M p_s(x) = 1 \pmod{\mu}, \quad (6.1.14)$$

$$p_s(x) = \int d\mu^s p_s(x, x_1, \dots, x_s), \quad s = 1, \dots, M.$$

Моделирование такого процесса осуществляется следующим образом.

Для $t = 0$ моделируем плотность $p^0(x)$, находим x_0 — координату “родившейся” частицы.

Если в момент времени t имеются частицы в точках $x(t, 1), \dots, x(t, k)$, то для каждой из них вычисляются $p_s(x(t, j))$, $s = 1, \dots, M$; $j = 1, \dots, k$; моделируется дискретное распределение $p_0(x(t, j)), \dots, p_M(x(t, j))$ и находится s . При $s = 0$ потомков нет — частица погибает. При $s > 0$ моделируем плотность $p_s(x(t, j), x_1, \dots, x_s)/p_s(x(t, j))$ — рождается s частиц. Таким образом, реализации процесса соответствует дерево. Если использовать более подробную запись для функционала $\mathfrak{J}(\gamma_n)$ на дереве, связанном с уравнением (6.1.11), а именно

$$\mathfrak{J}(\gamma_n) = \mathfrak{J}(\gamma_n, f, k_s, s = 1, \dots, M),$$

то для плотности вероятности $p(\gamma_n)$ дерева с не более чем n поколениями, порождаемого реализациями описанного выше процесса, будем иметь

$$p(\gamma_n) = \mathfrak{J}(\gamma_n, p_0, p_1, \dots, s = 1, \dots, M) \cdot p^0(x_0).$$

При фиксированном x_0 обозначим $p(\gamma_n) = p(\gamma_n, x_0)$. Из равенства (6.1.13) видим, что

$$\psi_n(x_0) = \sum_{\gamma_n} \int p(\gamma_n, x_0) d\mu^{|\gamma_n|}$$

удовлетворяет равенству

$$\psi_n(x_0) = p_0(x_0) + \sum_{s=1}^M \int d\mu^s p(x_0, x_1, \dots, x_s) \prod_{j=1}^s \psi_{n-1}(x_j), \quad (6.1.15)$$

и если предел последовательности $\psi_n(x_0)$ при $n \rightarrow \infty$ существует и $\lim_{n \rightarrow \infty} \psi_n(x_0) = \psi(x_0)$, то этот предел есть итерационное решение уравнения

$$\psi(x) = p_0(x) + \sum_{s=1}^M \int d\mu^s p(x_0, x_1, \dots, x_s) \prod_{j=1}^s \psi(x_j). \quad (6.1.16)$$

Легко видеть, что $\psi(x_0) \equiv 1$ удовлетворяет этому уравнению, — это следует из равенства (6.1.14).

Мы будем предполагать, что единица является итерационным решением уравнения (6.1.16), если сходится процесс (6.1.15). Легко показать, что он сходится по крайней мере при $p^0(x)$ близком к единице, то есть когда вероятность гибели частицы достаточно велика.

В этих условиях справедливо утверждение, что описанный нами ветвящийся процесс индуцирует вероятностную меру на множестве своих траекторий (деревьев).

Теперь по аналогии с линейным случаем построим алгоритм метода Монте-Карло для решения уравнений с полиномиальной нелинейностью. Относительно уравнения будем предполагать, что выполнено мажорантное условие: сходится метод последовательных приближений

$$\bar{\varphi}_n(x) = |f(x)| \sum_{s=1}^M \int d\mu^s |k_s(x, x_1, \dots, x_s)| \prod_{j=1}^s \bar{\varphi}_{n-1}(x_j), \quad \bar{\varphi}_0 = |f| \quad (6.1.17)$$

(для наших целей достаточно слабой сходимости, то есть существования предела $\lim_{n \rightarrow \infty} (h, \overline{\varphi_n})$ для всех h из некоторого множества функций $\mathcal{H}$). При этом предположении очевидно также, что для всех h из $\mathcal{H}$ существует предел последовательности (h, φ_n) при $n \rightarrow \infty$, где

$$\varphi_n = f + \sum_{s=1}^M \mathcal{K}_s \varphi_{n-1}^s, \quad \varphi_0 = f. \quad (6.1.18)$$

Здесь $\mathcal{K}_s$ — оператор с ядром k_s , а $\varphi^s = \varphi^s(x_1, \dots, x_s) = \prod_{j=1}^s \varphi(x_j)$. Переходные плотности ветвящегося процесса $p_j(x, x_1, \dots, x_s)$ будем выбирать удовлетворяющими условиям согласования:

- $p^0(x) > 0$ для всех $(\text{mod } \mu)$ x , для которых $h(x) \neq 0$;
- $p_0(x) > 0$ для всех $(\text{mod } \mu)$ x , для которых $f(x) \neq 0$;
- $p(x, x_1, \dots, x_s) > 0$ для всех $(\text{mod } \mu^s)$ векторов $(x, x_1, \dots, x_s)$, для которых $k_s(x, x_1, \dots, x_s) \neq 0$.

Обозначим через $Q(\gamma_n, h, f, k_s, s = 1, \dots, M)$ следующую функцию γ_n :

$$Q(\gamma_n, h, f, k_s, s = 1, \dots, M) = h(x_0) \mathfrak{J}(\gamma_n, f, k_s, s = 1, \dots, M).$$

Докажем теорему.

Теорема 6.1. Пусть γ_n — траектория случайного процесса, удовлетворяющего условиям согласования (6.1.19). Тогда случайная величина

$$\zeta(\gamma_n) = \frac{Q(\gamma_n, h, f, k_s, s = 1, \dots, M)}{Q(\gamma_n, p^0, p_0, p_s, s = 1, \dots, M)} \quad (6.1.20)$$

является несмещенной оценкой функционала (h, φ_n) , где φ_n удовлетворяет (6.1.18).

Доказательство. Для $E\zeta(\gamma_n)$ имеем выражение

$$\begin{aligned} E\zeta(\gamma_n) &= \int \sum_{\gamma_n} d\mu^{|\gamma_n|} \frac{Q(\gamma_n, h, f, k_s, s = 1, \dots, M)}{Q(\gamma_n, p_1^0, p_0, p_s, s = 1, \dots, M)} \times \\ &\quad \times Q(\gamma_n, p^0, p_0, p_s, s = 1, \dots, M). \end{aligned}$$

Условия согласования дают возможность сократить знаменатель дроби, а мажорантное условие позволяет менять местами суммирование и интегрирование. Поэтому

$$E\zeta(\gamma_n) = \sum_{\gamma_n} \int d\mu^{|\gamma_n|} Q(\gamma_n, h, f, k_s, s = 1, \dots, M) = (h, \varphi_n), \quad h \in \mathcal{H}. \quad (6.1.21)$$

□

Устремив n к бесконечности (это можно сделать, так как выполнено мажорантное условие), мы можем рассматривать деревья γ_n с любым числом поколений. Заметим, что из существования и единственности итерационного решения $\psi(x) \equiv 1$ уравнения (6.1.16) следует, что почти все траектории γ_n имеют конечное число поколений. Мажорантное условие позволяет перейти к пределу в (6.1.21) при $n \rightarrow \infty$, и мы получаем

Следствие. Имеет место равенство

$$E\dot{\zeta}(\gamma) = (h, \varphi), \quad h \in \mathcal{H}, \quad (6.1.22)$$

где φ — решение уравнения $\varphi = \sum_{s=1}^M \mathcal{K}_s \varphi^s + f$. $\square$

Легко видеть также, что, как и в линейном случае, условия согласования и мажорантные условия являются необходимыми и достаточными для выполнения равенства (6.1.22).

Практически реализация алгоритма осуществляется либо методом поколений, либо лексикографическим. Последний реализуется с помощью приема, известного под названием “магазин”. Отводится некая область памяти R , пропорциональная максимальному количеству M поколений в дереве. (Фактически мы считаем, что деревья с числом поколений, превышающим некоторое M , столь маловероятны, что их вкладом можно пренебречь.)

Далее выполняется следующая линейная процедура. Полагаем $V_0 = 1$; линейный массив R заполнен нулями.

1. $t = 0$. Находим x_0 , полагаем $V_1 = V_0 \cdot \frac{h_0}{p_0}$. Если далее частица погибает, то вычисляется $s = V_1 \cdot \frac{f_0}{g_0}$ и процесс оканчивается.
2. $t > 0$. Вычисляем s . При $s > 0$ находим $x_1^s(t), \dots, x_s^s(t)$. Числа $x_2^s(t), \dots, x_s^s(t)$ отсылаем в начало массива R , сдвигая имеющиеся там элементы.
3. Вычисляем $V_t = V_{t-1} \cdot \frac{k_s(x^s(t-1), x_2^s(t), \dots, x_s^s(t))}{p_s(x^{s-1}(t-1), x_1^s(t), \dots, x_s^s(t))}$ и возвращаемся к пункту 2.
4. Если $s = 0$, то $V_t = V_{t-1} \cdot \frac{f(x^s(t-1))}{g(x^{s-1}(t-1))}$ и оканчивается данная ветвь дерева. Здесь $x^s(t-1)$ — точка дерева, предшествующая координатам $x_j^s(t), j = 1, \dots, s$.

Если массив R пуст, то $\xi = \xi_t$ и вычисления заканчиваются. В противном случае берется первое x из R , для него разыгрывается s , полагается $x = x^s(t-1)$ и осуществляется переход к 2.

Метод поколений отличается лишь тем, что массив R заполняется координатами частиц, относящихся к поколению $t - 1$. Они обрабатываются по формулам блоками 2 и 3 описанного алгоритма. R заполняется координатами $x(t)$ появившихся вновь частиц. Если для очередного t массив пуст, то вычисления заканчиваются.

Дисперсия

Для оценки $\zeta(\gamma)$ легко получить выражение, аналогичное линейному случаю. Если величина $D\zeta(\gamma)$ конечна, то

$$\begin{aligned} D\zeta(\gamma) &= E\zeta^2(\gamma) - (E\zeta(\gamma))^2 = \\ &= \sum_{\gamma} \int d\mu^{|\gamma|} \frac{Q^2(\gamma, h, f, k_s, s = 1, \dots, M)}{Q(\gamma, p^0, p_0, p_s, s = 1, \dots, M)} - (\varphi, h)^2 = \\ &= \left(\frac{h^2}{p^0}, \psi\right) - (\varphi, h)^2, \end{aligned} \quad (6.1.23)$$

где ψ есть решение уравнения

$$\psi(x) = \frac{f^2(x)}{g(x)} + \sum_{s=1}^M \frac{k_s^2(x, x_1, \dots, x_s)}{p_s(x, x_1, \dots, x_s)} \prod_{j=1}^s \psi(x_j).$$

Также аналогично линейному случаю для полученной оценки строится оптимальная цепь Маркова, для которой в случае неотрицательных h, f и k_j , $j = 1, \dots, M$, дисперсия обращается в нуль. Для этой цепи

$$\begin{aligned} p^0 &= \frac{|h|\bar{\varphi}}{(|h|, \bar{\varphi})}, \quad p_s(x, x_1, \dots, x_s) = \frac{|k_s(x_1, \dots, x_s)|}{\bar{\varphi}(x)} \prod_{j=1}^s \bar{\varphi}(x_j), \\ g(x) &= \frac{|f(x)|}{\bar{\varphi}(x)}. \end{aligned}$$

Доказательство легко осуществить путем прямых выкладок по аналогии с линейной задачей.

Если мы используем описанный ветвящийся процесс, то построение других, и особенно двойственных, оценок для полиномиальной нелинейности, к сожалению, оказывается не столь простым, как для линейного случая. Далее изложим другие подходы, которые основаны на линеаризации задачи и способствуют дальнейшим обобщениям.

6.2. Уравнения с полиномиальной нелинейностью. Линеаризация

Формально интегральное уравнение

$$\varphi(x) = f(x) + \sum_{s=1}^M \int d\mu^s k_s(x, x_1, \dots, x_s) \prod_{j=1}^s \varphi(x_j) \quad (6.2.1)$$

можно свести к бесконечной системе линейных уравнений возрастающей размерности. Для этого достаточно обозначить $\psi_s(x_1, \dots, x_s) = \prod_{j=1}^s \varphi(x_j)$. Уравнение становится линейным относительно M неизвестных функций ψ_s . Домножая (6.2.1) на $\varphi(x_{M+1}), \varphi(x_{M+2}), \dots$, получим

$$\begin{aligned} \psi_2(x, x_{M+1}) &= \psi_1(x_{M+1})f(x) + \\ &+ \sum_{s=1}^M \int d\mu^s k_s(x, x_1, \dots, x_s) \psi_{s+1}(x_1, \dots, x_s, x_{M+1}), \\ \psi_3(x, x_{M+1}, x_{M+2}) &= \psi_2(x_{M+1}, x_{M+2})f(x) + \\ &+ \sum_{s=1}^M \int d\mu^s k_s(x, x_1, \dots, x_s) \psi_{s+2}(x_1, \dots, x_s, x_{M+1}, x_{M+2}) \end{aligned} \quad (6.2.2)$$

и т. д.

Такая линеаризация позволяет формально использовать рассмотренную в гл. 4 схему Неймана–Улама. Имеется весьма прозрачная связь с ветвящимся процессом предыдущего раздела, которую неформально можно описать следующим образом. Уравнение $\varphi(x) = \psi_1(x)$ “отвечает” за распределение одной частицы (объекта, особи), $f(x)$ ассоциируется с ее гибелью, а k_s — с появлением s новых частиц (потомков) (см. условия согласования). Положение s частиц описывается функцией ψ_s . Легко записать “большой” матричный оператор и применить к нему схему Неймана–Улама. Это позволяет строить оценку “по столкновениям” — аналог оценки $\mathfrak{J}_c^*$ в линейном случае.

Не развивая детально эту общую схему, остановимся на некоторых частных примерах ее применения. Сравнивая наш случай с линейным, заметим, что конструкция двойственного оператора осуществляется не столь просто. Условия согласования требуют введения в рассмотрение процесса, обратного ветвящемуся, — столкновительного. Он характерен тем, что за единицу времени s “частиц” могут превратиться

в одну. Непростые проблемы, относящиеся к столкновительным процессам, мы обсудим на частных примерах.

Рассмотрим сначала квадратное уравнение вида $x = ax^2 + bx + c$ в предположении, что сходится итерационный процесс:

$$x_n = |a|x_{n-1}^2 + |b|x_{n-1} + |c|, \quad x_0 = |c|. \quad (6.2.3)$$

Ветвящийся процесс здесь определяется распределением вероятностей p_2, p_1, p_0 — появления у некоторой “частицы” двух, одного или 0 потомков. Деревья строятся в соответствии с общими правилами, описанными в предыдущем разделе. Это, видимо, самый сложный из известных методов решения квадратного уравнения.

Цепочка уравнений (6.2.2) в этом случае имеет вид

$$ay_{k+1} + (b-1)y_k + cy_{k-1} = 0, \quad y_0 = 1, \quad k = 1, 2, \dots \quad (6.2.4)$$

Она получена заменой $y_k = x^k$. Это система разностных уравнений с постоянными коэффициентами. Исходное квадратное уравнение — характеристическое уравнение системы (6.2.4).

Если вернуться к представлению

$$y_k = ay_{k+1} + by_k + cy_{k-1}, \quad y_0 = 1, \quad k = 1, 2, \dots, \quad (6.2.5)$$

то легко видеть, что система характеризуется трехдиагональной матрицей. Ее решение методом Монте-Карло (схема Неймана–Улама) сводится к случайным блужданиям по решетке (на один шаг вправо или влево). Справа — поглощающая граница. Если при достаточно большом N положить $y_N = 0$, то получим поглощающую границу и слева. Приближенное решение будет при этом иметь систематическую погрешность. Очевидно, выбор сравнительно большого p_2 (вероятность перехода влево) может при $N \rightarrow \infty$ сделать задачу практически нерешаемой.

Здесь легко построить двойственную схему. Транспонируя бесконечную матрицу системы (6.2.5), имеем

$$y_k^* = cy_{k+1}^* + by_k^* + ay_{k-1}^*. \quad (6.2.6)$$

Не составит труда проверить, что выполняется соотношение

$$cy_i = a/y_i^*, \quad i = 1, 2, \dots,$$

которому удовлетворяет решение квадратного уравнения, полученного с помощью явной формулы.

Таким образом, линеаризация, основанная на домножении уравнения на произведения искомым функций, по крайней мере формально позволяет рассматривать то же семейство несмещенных оценок функционалов, что и в линейном случае. Далее мы обсудим аналогии с линейным случаем на примере системы алгебраических уравнений с квадратичной нелинейностью.

Рассмотрим систему уравнений вида

$$x_i = f_i + \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n \sum_{k=1}^n b_{ijk}x_jx_k, \quad i = 1, \dots, n, \quad (6.2.7)$$

которая является частным случаем уравнения (6.2.1). При поиске итерационного решения (6.2.7) предполагается сходимость мажорантного процесса

$$\bar{x}_i^{m+1} = |f_i| + \sum_{j=1}^n |a_{ij}| \bar{x}_j^m + \sum_{j=1}^n \sum_{k=1}^n |b_{ijk}| \bar{x}_j^m \bar{x}_k^m, \quad (6.2.8)$$

$$\bar{x}_i^0 = 0; \quad i = 1, \dots, n; \quad m = 0, 1, \dots$$

Домножая уравнение (6.2.7) на произведение

$$x_{i_1} \dots x_{i_l}, \quad i_1, \dots, i_l = 1, \dots, n,$$

и обозначая $U_l(i_1, \dots, i_l) = \prod_{k=1}^l x_{i_k}$, получаем бесконечную цепочку линейных соотношений

$$\begin{aligned} U_l(i_1, \dots, i_l) &= f_{i_1} U_{l-1}(i_2, \dots, i_l) + \sum_{j=1}^n a_{i_1,j} U_l(j, i_2, \dots, i_l) + \\ &+ \sum_{j=1}^n \sum_{k=1}^n b_{i_1,j,k} U_{l+1}(j, k, i_2, \dots, i_l), \quad (6.2.9) \\ i_1, \dots, i_l &= 1, 2, \dots, n, \quad l = 1, 2, \dots, \quad U_0 = 1. \end{aligned}$$

Таблица 6.1 содержит часть матрицы системы уравнений (6.2.9), соответствующей случаю $n = 2$. В таблице указаны номера уравнений (можно считать, что номер записывается в двоичной системе счисления с цифрами 1 и 2; это удобно для дальнейшего понимания). В аналогичном формате выписаны номера переменных.

Таблица 6.1.

Номера переменных уравнений	1	2	11	12	21	22	111	112	121	122	211	212	221	222
1	a_{11}	a_{21}	b_{111}	b_{112}	b_{121}	b_{122}	0	0	0	0	0	0	0	0
2	a_{12}	a_{22}	b_{211}	b_{212}	b_{221}	b_{222}	0	0	0	0	0	0	0	0
11	f_1	0	a_{11}	0	a_{12}	0	b_{111}	0	b_{112}	0	b_{121}	0	b_{122}	0
12	0	f_1	0	a_{11}	0	a_{12}	0	b_{111}	0	b_{112}	0	b_{121}	0	b_{122}
21	f_2	0	a_{21}	0	a_{22}	0	b_{211}	0	b_{212}	0	b_{221}	0	b_{222}	0
22	0	f_2	0	a_{21}	0	a_{22}	0	b_{211}	0	b_{212}	0	b_{221}	0	b_{222}
111	0	0	f_1	0	0	0	a_{11}	0	0	0	a_{12}	0	0	0
112	0	0	0	f_1	0	0	0	a_{11}	0	0	0	a_{12}	0	0
121	0	0	0	0	f_1	0	0	0	a_{11}	0	0	0	a_{12}	0
122	0	0	0	0	0	f_1	0	0	0	a_{11}	0	0	0	a_{12}
211	0	0	f_2	0	0	0	a_{21}	0	0	0	a_{22}	0	0	0
212	0	0	0	f_2	0	0	0	a_{21}	0	0	0	a_{22}	0	0
221	0	0	0	0	f_2	0	0	0	a_{21}	0	0	0	a_{22}	0
222	0	0	0	0	0	f_2	0	0	0	a_{21}	0	0	0	a_{22}

Продолжение матрицы легко себе представить: оно состоит из блоков размерности 4 и 8 вместо 2 и 4. Случай $n > 2$ не имеет принципиальных особенностей.

Условия согласования, аналогичные (6.1.19), предполагают, что соответствующая системе (6.2.8) цепь Маркова должна иметь $p_i^{(0)}$ на месте f_i ($p_i^{(0)} > 0$, если $f_i \neq 0$), $p_{i,j}^{(1)}$ на месте a_{ij} ($p_{i,j}^{(1)} > 0$, если $a_{ij} \neq 0$) и $p_{i,j,k}^{(2)}$ на месте $b_{i,j,k}$ ($p_{i,j,k}^{(2)} > 0$, если $b_{i,j,k} \neq 0$). Остальные элементы переходной матрицы могут быть нулями (что мы и полагаем). При этом

$$p_i^{(0)} + \sum_{j=1}^n p_{i,j}^{(1)} + \sum_{j=1}^n \sum_{k=1}^n p_{i,j,k}^{(2)} = 1, \quad i = 1, \dots, n. \quad (6.2.10)$$

Заметим также, что правая часть системы имеет вид $(f_1, f_2, \dots, f_n, 0, \dots)^T$, что определяет структуру распределения вероятности поглощения — $(p_1^0, p_2^0, \dots, p_n^0, 0, \dots, 0)^T$.

Вернемся к табл. 6.1, мысленно заменив в ней f на p^0 , a на $p^{(1)}$ и b на $p^{(2)}$ (с соответствующими индексами, разумеется) и рассмотрим цепь Маркова, имеющую переходную матрицу такой структуры. Рассуждение легко распространить на произвольное n .

Как и ранее, предполагаем, что время дискретно. При $t = 0$ имеем частицу в состоянии i_0 с вероятностью π_{i_0} . Если она находится в состоянии i_0 , то далее с вероятностью $p_{i_0}^{(0)}$ — погибает, с вероятностью $p_{i_0,j}^{(1)}$ — переходит в состояние j , а с вероятностью $p_{i_0,j,k}^{(2)}$ — в состояние (j, k) (строки с номерами 11, 12, 21, 22 в табл. 6.1). В последнем случае можно считать, что появились 2 частицы в состояниях j и k соответственно.

Если появились две частицы, то первая подматрица, которая для случая $n = 2$ имеет вид

$$\begin{pmatrix} p_1^0 & 0 \\ 0 & p_1^0 \\ p_2^0 & 0 \\ 0 & p_2^0 \end{pmatrix},$$

включает в себя вероятности состояний двух частиц, ответственные за то, что одна из них погибнет; второй блок описывает изменение конфигурации состояний обеих частиц (из (i, j) в (k, l) ; $i, j, k, l = 1, 2$), а следующие 2 блока в табл. 6.1, имеющие размерность 4×4 , отвечают за образование третьей частицы и за конфигурацию состояний трех частиц соответственно.

Нетрудно проследить продолжение процесса. Понятно, что мы имеем дело с ветвящимся процессом, только моделирование и обработка его траекторий отличаются от описанных в предыдущем разделе. Преимущество такого подхода состоит в прозрачности переноса на нелинейный случай методов построения оценок, развитых для линейного случая.

Используя условия согласования, мы также можем интерпретировать цепь Маркова для решения системы в терминах поведения частиц.

В нашей задаче речь идет о поведении ансамбля из l частиц. Процесс с дискретным временем $t = 0, 1, 2, \dots$, переходная матрица которого согласована с матрицей вида 6.1, можно описать следующим образом (эту матрицу легко представить в обобщенном виде).

Опишем подробно алгоритм вычисления оценок.

При $t = 0$ имеется одна частица, положение которой i_0 определяется начальным распределением $\vec{\pi} = (\pi_0, \dots, \pi_n)$.

Если при $t > 0$ существует $l > 1$ частиц, то одна из них выбирается случайно с равными вероятностями и для нее в момент времени $t + 1$ имеются следующие возможности:

- с вероятностью $p_{i_t}^{(0)}$ частица погибает,
- с вероятностью $p_{i_t, i_{t+1}}^{(1)}$ — переходит в состояние с номером i_{t+1} ,
- с вероятностью $p_{i_t, i_{t+1}, i'_{t+1}}^{(2)}$ появляются 2 частицы в состояниях с номерами i_{t+1} и i'_{t+1} соответственно.

Остальные частицы в момент времени $t+1$ не меняют своих состояний.

Статистические веса Q_t вычисляются одновременно с моделированием поведения ансамбля частиц в соответствии со следующими формулами для каждой из указанных возможностей:

$$\begin{aligned}
 (1) \quad Q_0 &= h_{i_0}/\pi_{i_0}; \quad Q_{t+1} = Q_t f_{i_t}/p_{i_t}^0, \quad l > 1, \\
 (2) \quad Q_{t+1} &= Q_t a_{i_t, i_{t+1}}/p_{i_t, i_{t+1}}^1, \\
 (3) \quad Q_{t+1} &= Q_t b_{i_t, i_{t+1}, i'_{t+1}}/p_{i_t, i_{t+1}, i'_{t+1}}^2.
 \end{aligned} \tag{6.2.11}$$

Если $l = 1$, то в варианте (1) траектория обрывается ($t = \tau$) с вероятностью q_{i_τ} и величина $\xi_\tau^{(1)} = Q_\tau f_{i_\tau}/q_{i_\tau}$ является несмещенной оценкой скалярного произведения (H, X) , где X — итерационное решение системы (6.2.7). Назовем $\xi_\tau^{(1)}$ оценкой по поглощению аналогично линейному случаю.

Конструкция аналога оценки по столкновениям также сходна с линейным случаем, однако имеет некоторые специфические особенности, связанные с видом правой части системы (6.2.9). Траектория может оборваться лишь при $l=1$ (условие согласования). Если $t_1, \dots, t_s$ — моменты времени, когда моделируемый ансамбль состоит только из одной частицы, то оценка по столкновениям примет вид:

$$\mathfrak{J}_a(\omega_\tau) = \sum_{i=1}^s Q_{t_i} f_{t_i}. \quad (6.2.12)$$

Теперь можно (по крайней мере формально) описать цепь Маркова, переходная матрица которой согласована с транспонированной матрицей системы уравнений (6.2.9). Правая часть двойственной системы есть $(h_1, \dots, h_n, 0, 0, \dots)$.

Условия согласования

$$\begin{aligned} p_{i,j,k}^2 &> 0, & \text{если } b_{i,j,k} &\neq 0; \\ p_{i,j}^1 &> 0, & \text{если } a_{i,j} &\neq 0; \\ p_i^0 &> 0, & \text{если } f_i &\neq 0, \end{aligned} \quad (6.2.13)$$

и структура процесса определяются структурой матрицы $\mathcal{P}'$, которая имеет ненулевые элементы на местах ненулевых элементов матрицы системы (6.2.8). Первые столбцы и строки $\mathcal{P}'$ приведены в таблице 6.2. Легко видеть, что $\mathcal{P}'$ имеет блочную структуру и размерность блоков на главной диагонали — n^l , где l также соответствует количеству частиц в ансамбле. Каждая строка содержит $p_{i,j,k}^2, p_{i,j}^1, p_i^0$, $i = 1, 2, \dots, n$. Первые n строк не содержат $p_{i,j,k}^2$.

В соответствии со сказанным должны выполняться следующие условия нормировки:

$$\begin{aligned} \sum_{i=1}^n (p_{i,j,k}^2 + p_{i,j}^1 + p_i^0) &= 1, \quad j, k = 1, \dots, n, \\ \sum_{i=1}^n p_{i,j,k}^2 &= p_j^2 \quad \text{не зависит от } k, \\ \sum_{i=1}^n p_{i,j}^1 &= p_j^1, \quad p_j^2 + p_j^1 \quad \text{не зависит от } j. \end{aligned} \quad (6.2.14)$$

Процесс можно интерпретировать следующим образом. При $t = 0$ появляется одна частица в состоянии со случайным номером, распределение которой согласовано с вектором H .

Таблица 6.2.

Номера переменных уравнений	1	2	11	12	21	22	111	112	121	122	211	212	221	222
1	a_{11}	a_{21}	f_1	0	f_2	0	0	0	0	0	0	0	0	0
2	a_{12}	a_{22}	0	f_1	0	f_2	0	0	0	0	0	0	0	0
11	b_{111}	b_{211}	a_{11}	0	a_{21}	0	f_1	0	0	0	f_2	0	0	0
12	b_{112}	b_{212}	0	a_{11}	0	a_{21}	0	f_1	0	0	0	f_2	0	0
21	b_{121}	b_{221}	a_{12}	0	a_{22}	0	0	0	f_1	0	0	0	f_2	0
22	b_{122}	b_{222}	0	a_{12}	0	a_{22}	0	0	0	f_1	0	0	0	f_2
111	0	0	b_{111}	0	b_{211}	0	a_{11}	0	0	0	a_{21}	0	0	0
112	0	0	0	b_{111}	0	b_{211}	0	a_{11}	0	0	0	a_{21}	0	0
121	0	0	b_{112}	0	b_{212}	0	0	0	a_{11}	0	0	0	a_{21}	0
122	0	0	0	b_{112}	0	b_{212}	0	0	0	a_{11}	0	0	0	a_{21}
211	0	0	b_{121}	0	b_{221}	0	a_{12}	0	0	0	a_{22}	0	0	0
212	0	0	0	b_{121}	0	b_{221}	0	a_{12}	0	0	0	a_{22}	0	0
221	0	0	b_{122}	0	b_{222}	0	0	0	a_{12}	0	0	0	a_{22}	0
222	0	0	0	b_{122}	0	b_{222}	0	0	0	a_{12}	0	0	0	a_{22}

Если в момент времени t имеется l частиц, то

- с вероятностью p_i^0 появляется новая частица в состоянии с номером i — число частиц становится $l + 1$;
- с вероятностью $p_{i,j}^1$ случайно выбранная частица меняет номер своего состояния на i , если она находилась в состоянии с номером j ;
- при $l > 1$ две частицы, находящиеся в состояниях с номерами j и k соответственно, сольются и образуют одну частицу в состоянии с номером i . При $l = 1$ единственная частица погибает в состоянии i с вероятностью p_i^2 .

Процедура вычисления статистических весов в процессе моделирования выглядит следующим образом.

При $t = 0$ рождается одна частица в состоянии i_0 и вычисляется $Q_0 = f_{i_0}/p_{i_0}^0$. При $t > 0$ в первом случае имеем $Q_{t+1} = Q_t \cdot f_{i_t} p_{i_t}^0$; во втором — $Q_{t+1} = Q_t \cdot a_{i_{t+1}, i_t} / p'_{i_{t+1}, i_t}$; в третьем при $l > 1$ — $Q_{t+1} = Q_t b_{i_{t+1}, i_t, i_t^1} / p_{i_{t+1}, i_t, i_t^1}^2$, а при $l = 1$ процесс обрывается с вероятностью $p_{i_t}^2$ ($t = \tau$) и вычисляется оценка по поглощению

$$\mathfrak{J}_a^*(\omega_\tau) = Q_\tau \frac{h_{i_\tau}}{p_{i_\tau}^2}. \quad (6.2.17)$$

Оценка по столкновениям вычисляется по вкладам Q_{t_i} , где t_i — моменты времени, когда ансамбль состоял из одной частицы ($l = 1$)

$$\mathfrak{J}_c^*(\omega_\tau) = \sum_{i=1}^s Q_{t_i} h_{t_i}. \quad (6.2.18)$$

Следует отметить, что использование двойственности в нелинейном случае требует, как правило, дополнительных исследований. Наша теория безупречна, лишь если система (6.2.9) приближенно заменяется конечной системой. Переход к пределу нуждается в обосновании. Отметим также, что для двойственной системы не обязательно существует такой полиномиальный оператор, для которого эта система принимает вид (6.2.8).

Обобщения для интегральных операторов в квадратичном случае и для нелинейностей более высокой степени не составляют труда, хотя соответствующие доказательства требуют громоздких выкладок. Кроме того, мажорантные условия с ростом степени становятся все

более ограниченными. По-видимому, более перспективно использование рекуррентных процедур, аналогичных описанным в гл. 5, но исследования в этой области находятся в начальной стадии. Возникающим здесь проблемам посвящен п. 6.3.

6.3. Стохастическая устойчивость. Нелинейный случай

6.3.1. Стохастический аналог метода Ньютона

Пусть $X = (x_1, \dots, x_n)$ — вектор неизвестных, $G(X) = (g_1(x), \dots, g_s(x))$ — вектор-функция, обладающая достаточной степенью гладкости,

$$G(X) = 0 \quad (6.3.1)$$

— система нелинейных уравнений, подлежащая решению. Более определенно ставится задача численного нахождения одного из корней системы. Корень предполагается простым, а система (6.3.1) — преобразованной в систему

$$X = F(X), \quad F = (f_1, \dots, f_n), \quad X = (x_1, \dots, x_n), \quad (6.3.2)$$

таким образом, что метод итераций

$$X_m = F(X_{m-1}), \quad m = 1, 2, \dots, \quad (6.3.3)$$

сходится к упомянутому корню при заданном начальном X_0 . Условия, налагаемые на F для сходимости (6.3.3), достаточно полно описаны в литературе (например, в [62]), и мы остановимся лишь на стохастических аспектах проблемы, предполагая, что производная Фреше оператора F в окрестности корня имеет норму меньше единицы и X_0 принадлежит этой окрестности. В данном случае производная F есть матрица $\|\partial f_i / \partial x_j\|_{i,j=1}^n$. Возможны различные способы рандомизации итерационного процесса. Мы ограничимся лишь рассмотрением некоторых итерационных процессов, включающих на каждом этапе решение системы линейных алгебраических уравнений методом Монте-Карло. К числу таких методов можно отнести метод Ньютона и его модификации. В этом случае $\tilde{F}$ конструируется следующим образом:

$$\tilde{F}(X) = X - G'^{-1}(X)G(X), \quad (6.3.4)$$

где $G'(X)$ есть матрица производных $\|\partial g_i / \partial x_j\|_{i,j=1}^n$. Метод имеет второй порядок сходимости, если система уравнений с матрицей $G'(X_n)$

и правой частью $G(X_n)$ при каждом n решается достаточно точно. Если n велико и для ее решения используется метод Монте-Карло с умеренным числом независимых испытаний, то имеется случайная погрешность, влияние которой должно быть исследовано. Пусть N обозначает число испытаний при каждом m . Тогда вместо равенства (6.3.3) мы будем иметь $\Xi_m = \tilde{F}(\Xi_{m-1})$, где $\Xi_m = X_m + \mathcal{E}_m$ и $\Xi_{m-1} = \hat{\Xi}_{m-1}(N)$ — среднее N независимых испытаний. Здесь $\tilde{F}(X) = X - (\tilde{G}'^{-1})(X)G(X)$. Волна в правой части, как и ранее, означает, что матрица G' обращается методом Монте-Карло.

Подробный анализ статистических свойств случайного итерационного процесса

$$\Xi_m = \tilde{F}(\Xi_{m-1}), \quad \Xi_0 = X_0, \quad (6.3.5)$$

связан с громоздкими выкладками. Обоснованием использования такого процесса для вычислений при некотором N_0 , $N \geq N_0$, может служить следующий известный в математической статистике факт.

Мы сошлемся на утверждение 9.3.1а на с. 271 книги Уилкса [77].

Утверждение. Пусть $\xi_{1,i}, \dots, \xi_{n,i}$, $i = 1, \dots, N$, — выборка из n -мерного нормального распределения с конечными средними значениями μ_l и положительно определенной ковариационной матрицей $\|\sigma_{k,l}\|$, $k, l = 1, \dots, n$. Пусть также $g(x_1, \dots, x_n)$ — функция, имеющая первые производные $\frac{\partial g}{\partial x_i} = g_i$, $i = 1, \dots, n$, во всех точках некоторой окрестности точки $(\mu_1, \dots, \mu_n)$, и $g_i^0 = g_i(\mu_1, \dots, \mu_n)$. Тогда если по крайней мере одна из величин g_i^0 не равна нулю, то $g(\bar{x}_1, \dots, \bar{x}_n)$ имеет для больших n асимптотическое распределение

$$\mathcal{N}(g(\mu_1, \dots, \mu_n), \frac{1}{N} \sum_{k,l=1}^n \sigma_{k,l} g_k^0 g_l^0) \quad (6.3.6)$$

($\bar{x}_i$ — среднее соответствующих выборочных значений, а $\mathcal{N}(a, \sigma^2)$ — нормальное распределение со средним a и дисперсией σ^2). $\square$

Замечание. Из этого утверждения вытекает, в частности, что моменты выше второго порядка случайной величины $g(\bar{x}_1, \dots, \bar{x}_n)$ имеют порядок малости выше чем $1/N$. $\square$

Ограничимся рассмотрением лишь модификации метода Ньютона, полагая $G' = C$ — матрицей, не зависящей от X . Из равенств

$$\Xi_m = \Xi_{m-1} - \tilde{C}^{-1}G(\Xi_{m-1}) \quad \text{и} \quad X_m = X_{m-1} - \tilde{C}^{-1}G(X_{m-1})$$

имеем для $\vec{\mathcal{E}}_m = \Xi_m - X_m$

$$\begin{aligned}\vec{\mathcal{E}}_m &= \vec{\mathcal{E}}_{m-1} - \tilde{C}^{-1}G(\Xi_{m-1}) + C^{-1}G(X_{m-1}) - \\ &\quad - C^{-1}G(\Xi_{m-1}) + C^{-1}G(X_{m-1}).\end{aligned}$$

Обозначая $\Delta = \tilde{C}^{-1} - C^{-1}$, $\partial G/\partial X = \|\partial g_l/\partial x_j\|_{j,l=1}^n$, получим

$$\begin{aligned}\vec{\mathcal{E}}_m &= \vec{\mathcal{E}}_{m-1} - \Delta \left(G(X_{m-1}) + \left(\frac{\partial G(X_{m-1})}{\partial X} \vec{\mathcal{E}}_{m-1} \right) - \right. \\ &\quad \left. - C^{-1} \left(\frac{\partial G(X_{m-1})}{\partial X} \vec{\mathcal{E}}_{m-1} \right) \right) + O\left(\frac{1}{\sqrt{N}}\right)\end{aligned}$$

или

$$\vec{\mathcal{E}}_m = \left(I - C^{-1} \frac{\partial G(X_{m-1})}{\partial X} - \Delta \frac{\partial G(X_{m-1})}{\partial X} \right) \vec{\mathcal{E}}_{m-1} - \Delta G(X_{m-1}) + O\left(\frac{1}{\sqrt{N}}\right). \quad (6.3.7)$$

Если учесть, что второй момент Δ имеет порядок $O(\frac{1}{N})$ (метод Монте-Карло), то видно, что $\vec{\mathcal{E}}_m$ является асимптотически несмещенной оценкой нуля. Связь ковариаций $\vec{\mathcal{E}}_m$ и $\vec{\mathcal{E}}_{m-1}$ теперь получить нетрудно: достаточно выразить $E(\vec{\mathcal{E}}_m \vec{\mathcal{E}}_m^T)$ через $E(\vec{\mathcal{E}}_{m-1} \vec{\mathcal{E}}_{m-1}^T)$, используя равенство (6.3.7), как это было сделано в линейном случае. Дальнейший анализ может привести к оценке нужного при каждом m числа испытаний N , для которого процесс будет стохастически устойчивым. Здесь важно отметить, что явление стохастической неустойчивости может иметь место и в нелинейном случае.

Метод Ньютона оказывается особенно простым для рассматривавшейся ранее квадратичной нелинейности. Действительно, в случае системы уравнений

$$x_i = f_i + \sum_{j=1}^n a_{i,j} x_j + \sum_{j=1}^n \sum_{k=1}^n b_{i,j,k} x_j x_k, \quad i = 1, \dots, n, \quad (6.3.8)$$

$$G'(X) = \|a_{i,j} - \delta_{i,j} + \sum_{k=1}^n (b_{i,k,j} x_k + b_{i,j,k} x_k)\|_{i,j=1}^n. \quad (6.3.9)$$

В нашем примере $G'(X)$ не зависит от X ; детерминированный метод здесь имеет первый порядок сходимости. В общем случае сходимость, как известно, квадратичная. На этом основании можно ожидать, что при большом N норма оператора, определяющего стохастическую устойчивость метода, будет разве лишь меньше в случае зависимости от X при пересчете $G'(X)$ на каждой итерации. Численные эксперименты подтверждают это.

Уравнения механики жидкости и газа (Больцмана, Навье–Стокса) имеют квадратичную нелинейность. Их численное решение во многих случаях сводится к решению системы вида (6.3.8). С другой стороны, широкое распространение имеют методы прямого моделирования (Бёрд), основанные на имитации поведения молекул. Имитационные методы можно рассматривать так же, как методы моделирования марковской цепи, включающие расчет движения молекул и характеристик их парных взаимодействий. Сопоставление методов имитационного моделирования и стохастических методов решения соответствующих уравнений представляет интерес как с точки зрения повышения эффективности моделирования, так и с точки зрения улучшения или даже возникновения новых детерминированных методов. Не входя в детали имитационного моделирования, которому посвящена обширная литература, отметим следующее.

Имитационное моделирование в газовой динамике рассматривает, грубо говоря, “превращение” в результате взаимодействия двух молекул в две молекулы с другими характеристиками (координатами, скоростями и т. п.).

Методы решения уравнений с квадратичной нелинейностью, описанные ранее, предполагают “превращение” двух частиц в одну (столкновительный процесс) или одной частицы в две (ветвящийся процесс). По-видимому, конструкция процесса “превращения” двух частиц в две приблизилась бы нас к методам имитации. Ниже приводится описание такого процесса.

6.3.2. Квадратичная нелинейность. Метод искусственного хаоса

Рассмотрим уравнение с квадратичной нелинейностью вида

$$\varphi(x) = f(x) + \int k_1(x, u)\varphi(u)\mu(du) + \int k_2(x, u, v)\varphi(u)\varphi(v)\mu(du)\mu(dv), \quad (6.3.10)$$

где μ — σ -конечная мера на $\mathfrak{X}$; f задана на носителе меры μ , k_1 — на носителе μ^2 , а k_2 — на носителе μ^3 соответственно; φ — искомая функция.

Пусть $y \in \mathfrak{X}$. Домножая уравнение (6.3.10) на $\varphi(y)$, получим

$$\begin{aligned} \varphi(x)\varphi(y) &= f(x)\varphi(y) + \int k_1(x, u)\varphi(u)\varphi(y)\mu(du) + \\ &+ \int k_2(x, u, v)\varphi(y)\varphi(u)\varphi(v)\mu(du)\mu(dv). \end{aligned}$$

Пусть теперь $\bar{\varphi}$ — решение уравнения (6.3.10). Тогда $\bar{\psi}(x, y) = \bar{\varphi}(x)\bar{\varphi}(y)$ есть решение уравнения

$$\begin{aligned} \psi(x, y) = & f(x)\bar{\varphi}(y) + \int k_1(x, u)\psi(u, y)\mu(du) + \\ & + \int k_2(x, u, v)\bar{\varphi}(y)\psi(u, v)\mu(du)\mu(dv). \end{aligned} \quad (6.3.11)$$

Имеем линейное уравнение в пространстве удвоенной размерности. Схема Неймана–Улама для его решения моделирует “превращение” двух частиц в две. Рассмотрим сначала детерминированный итерационный процесс.

Будем искать решение ψ уравнения (6.3.11) методом последовательных приближений вида

$$\begin{aligned} \psi^m(x, y) = & f(x)\varphi^{m-1}(y) + \int k_1(x, u)\psi^m(u, y)\mu(du) + \\ & + \int k_2(x, u, v)\varphi^{m-1}(y)\psi^m(u, v)\mu(du)\mu(dv), \end{aligned} \quad (6.3.12)$$

где $m = 1, 2, \dots$

Задавая φ_0 и способ вычисления φ^{m-1} при известной ψ^{m-1} , решаем при каждом m линейное уравнение относительно $\psi^m(x, y)$ — функции двух переменных из пространства $\mathfrak{X}$.

Целесообразно добиваться такого выбора функции φ^{m-1} , при котором ψ^{m-1} можно приближенно представить в виде произведения $\psi^m(x, y) = \varphi^m(x)\varphi^m(y)$, — в этом заключается основная идея методов искусственного хаоса.

Можно указать несколько способов расчета φ^{m-1} на основе ψ^{m-1} :

- (1) Если $\psi^{m-1}(x, y) \geq 0$, то полагаем

$$\varphi^{m-1}(x) = \sqrt{\psi^{m-1}(x, x)} \quad (6.3.13)$$

при каждом m .

- (2) В некоторых случаях ψ^{m-1} меняет знак. Тогда при условии

$$\int \bar{\varphi}(y)\mu(dy) > 0$$

МОЖНО ВЗЯТЬ

$$\varphi^{m-1}(x) = \frac{\int \psi^{m-1}(x, t)\mu(dt)}{\sqrt{\int \psi^{m-1}(s, t)\mu(ds)\mu(dt)}} \quad (6.3.14)$$

или

$$\varphi^{m-1}(x) = \frac{\int (\psi^{m-1}(x, t) + \psi^{m-1}(t, x)) \mu(dt)}{2\sqrt{\int \psi^{m-1}(s, t) \mu(ds) \mu(dt)}}. \quad (6.3.15)$$

Итерационный процесс сохраняет нелинейность, определяемую формулами (6.3.13) или (6.3.14)–(6.3.15). Вычисление производной Фреше нелинейного оператора, задающего этот процесс, представляет собой чисто техническую задачу. Эти вычисления выполнены в [39] и здесь не воспроизводятся. Отметим лишь следующее.

Операторы $\mathcal{K}_1$ и $\mathcal{K}_2$, участвующие в вычислениях, имеют ядро $k_1(x, u)\delta_y(v)$ и $k_2(x, u, v)\varphi(y)$ соответственно. Требуется выполнять обращение оператора

$$I - \mathcal{K}_1 - \mathcal{K}_2. \quad (6.3.16)$$

При использовании метода Монте-Карло это означает, что при каждом m должно быть удовлетворено условие

$$\|\overline{\mathcal{K}_1 + \mathcal{K}_2^m}\| < 1, \quad (6.3.17)$$

где оператор $\overline{\mathcal{K}_1 + \mathcal{K}_2^m}$ имеет ядро $|k_1(x, u)\delta_y(v) + k_2(x, u, v)\varphi^m(y)|$.

Условие (6.3.17) говорит о том, что метод, вообще говоря, удобен лишь для решения специального класса задач.

Рассмотрим более подробно его особенности, если мера μ дискретна. В этом случае из (6.3.8) описанным выше приемом получаем систему уравнений

$$y_{i,l}^m = f_i x_l^{m-1} + \sum_{j=1}^n a_{i,j} y_{j,l}^m + \sum_{j,k=1}^n b_{i,j,k} x_l^{m-1} y_{j,k}^{m-1}, \quad (6.3.18)$$

$$i, l = 1, \dots, n,$$

x_l^{m-1} и $y_{j,k}^{m-1}$ связаны между собой одним из соотношений: (6.3.13) или (6.3.14)–(6.3.15).

Система (6.3.18) имеет размерность n^2 . Замена системы n уравнений системой n^2 уравнений в детерминированном случае смысла не имеет. Однако использование метода Монте-Карло приближает по трудоемкости задачу решения системы (6.3.18) к методу Ньютона, так как при каждом m оценивается лишь n величин (x_l , $l = 1, \dots, n$) и матрица системы имеет блочную структуру, сходную со структурой соответствующей матрицы (6.3.9) метода Ньютона.

Таким образом, на примере простейшей квадратичной нелинейности можно видеть, что существует большое разнообразие итерационных методов решения задачи. Последний из рассмотренных — метод искусственного хаоса — характерен тем, что его целесообразно применять лишь в стохастическом варианте.

6.3.3. Искусственный хаос и метод Бёрда в газовой динамике

Представляет интерес подробное сравнение метода искусственного хаоса с известным методом Бёрда решения задач газовой динамики.

Соотношения (6.3.15) в нашем случае могут быть представлены в виде

$$x_l^{m-1} = \left(\sum_{i=1}^n y_{i,l}^{m-1} + \sum_{j=1}^n y_{l,j}^{m-1} \right) / \left(2 \sum_{i,j} y_{i,j}^{m-1} \right)^{1/2}$$

или

$$x_l^{m-1} = \sum_{i=1}^n y_{l,i}^{m-1} / \left(\sum_{i,j} y_{i,j}^{m-1} \right)^{1/2}, \quad (6.3.19)$$

(суммы в числителе (6.3.19) должны быть равны). При вычислении $y_{i,j}^m$ можно использовать вспомогательный (внутренний) итерационный процесс. При фиксированных m и x_l^{m-1} обозначим

$$F_{i,j} = x_j^{m-1} \left(f_i + \sum_{j,k} b_{i,j,k} y_{j,k}^m \right), \quad (6.3.20)$$

$$\mathcal{F}_p = (F_{1,p}, \dots, F_{n,p})^T$$

и будем решать систему

$$Y_p = AY_p + \mathcal{F}_p; \quad p = 1, 2, \dots, n; \quad \mathcal{F}_p = \mathcal{F}_p(Y_p) \quad (6.3.21)$$

методом итераций

$$Y_p^r = AY_p^r + \mathcal{F}_p(Y_p^{r-1}); \quad r = 1, 2, \dots$$

Y_p^0 полагаем равным Y_p , полученному на $m-1$ приближении внешнего итерационного процесса.

Заметим, что обсуждаются методы решения систем при большом n , при этом вычисление $\mathcal{F}_p(Y_p)$ может быть весьма трудоемким. В общем случае необходимо сложить $n(n+1)/2$ попарных произведений. В этих условиях целесообразна рандомизация, при которой сумма

очень большого числа слагаемых оценивается с помощью суммы меньшего числа случайно выбранных слагаемых. Пусть $\hat{\mathcal{F}}_p$ — несмещенная оценка $\mathcal{F}_p$ с конечной дисперсией. Далее поступим следующим образом:

1. Полагаем $r = 1$.
2. Выбираем случайный номер p , $1 \leq p \leq n$. Для каждого p вероятность его выбора положительна.
3. Вычисляем $\hat{\mathcal{F}}_p^r = \hat{\mathcal{F}}_p(Y_p^{r-1})$.
4. Пункты 2, 3 можно повторить N раз и вычислить среднее $\hat{\mathcal{F}}_{p,N}^r$ полученных значений $\hat{\mathcal{F}}_p^r$.
5. Система решается методом Монте-Карло. Предполагается $\rho(|A|) < 1$, $Y_p^r = AY_p^{r-1} + \hat{\mathcal{F}}_{p,N}^r$. Находится решение Ξ_p^r .
6. Переходим к пункту 1, полагая $r := r + 1$.

Процесс прекращается, когда решение установится. Критерий установления при данном m должен быть определен дополнительно. Несложные рассуждения показывают, что случайный выбор величин p (например, выбор их с равной вероятностью) может осуществляться так, чтобы выполнялось равенство

$$\mathbb{E}\Xi_p^r = (I - A)^{-1}\mathbb{E}\hat{\mathcal{F}}_{p,N}^r = \sum_p Y_p^{(m)}, \quad Y_p^{(m)} = (y_{p,1}^m, \dots, y_{p,n}^m) \quad (6.3.22)$$

в предположении, что вычисляемые оценки не зависят от оценок $\hat{\mathcal{F}}_{p,N}^r$. Это предположение выполняется во всех случаях, когда при вычислении оценок используются различные случайные числа.

Таким образом, строятся оценки сумм векторов $Y_p^{(m)}$, но именно компоненты этих сумм входят в (6.3.19) и определяют x_l^{m-1} . Алгоритм описан в общих чертах. Он позволяет не увеличивать размерность линейной системы.

Читатели, знакомые с методом Бёрда в газовой динамике (см. [2], [25]), сразу же увидят полную аналогию между ним и описанным алгоритмом. У Бёрда обращение матрицы A — это движение молекул, а выбор Y_p^{r-1} — это выбор пары молекул, которые сталкиваются, — $n(n+1)/2$ комбинаций. Целесообразность использования таких алгоритмов в других задачах требует отдельного рассмотрения. Наиболее интересны такого рода исследования в связи с методами Монте-Карло для решения уравнений Навье–Стокса.

6.3.4. Заключительные замечания. Параллелизм

Укажем еще один очень простой метод линеаризации, требующий решения систем линейных уравнений. Для дискретного случая это последовательность итераций

$$x_i^m = f_i + \sum_j a_{i,j} x_j^m + \sum_{j,k} b_{i,j,k} x_j^m x_k^{m-1}.$$

Здесь на каждой итерации также решается система линейных уравнений, похожая на системы методов Ньютона и искусственного хаоса. Во всех трех случаях возникает проблема стохастической устойчивости, которую мы исследовали для случая модифицированного метода Ньютона.

Далее можно сравнить особенности методов, основанных на моделировании ветвящихся процессов, и трех других, связанных с решением систем линейных алгебраических уравнений на каждой итерации. Отметим в первую очередь, что каждый метод связан с цепями Маркова, которые должны удовлетворять условиям согласования, вообще говоря, различным для каждого из них. То же относится и к мажорантным условиям.

С точки зрения параллелизма, методы, связанные с ветвящимися процессами, дают несмещенные оценки решения и обладают естественным параллелизмом. Их недостатком являются жесткие мажорантные условия, однако применимы методы последовательного оценивания итераций, также не имеющие смещения, но требующие синхронизации (крупнозернистый параллелизм).

Методы, связанные с системами линейных алгебраических уравнений, дают асимптотически несмещенные оценки. При осуществлении вычислений на нескольких процессорах необходимо обеспечить как стохастическую устойчивость, так и достаточно малое смещение. Оказывается, что при большом числе процессоров и, следовательно, малом числе испытаний на каждой итерации:

- метод Ньютона имеет линейный порядок сходимости (как и два других);
- для стохастической устойчивости методов искусственного хаоса и прямой линеаризации может потребоваться меньше испытаний “внутри итерации”, чем для метода Ньютона.

Соответствующие примеры решения простых алгебраических систем приводятся в приложении 4.

Ниже мы кратко остановимся на уже опробованных методах решения уравнений Навье–Стокса.

6.4. О решении разностного аналога уравнений Навье–Стокса

Исходные уравнения в инерциальной декартовой системе координат имеют вид

$$\frac{\partial \vec{U}}{\partial t} - \nu \Delta \vec{U} + \sum_{i=1}^3 u_i \frac{\partial \vec{U}}{\partial x_i} + \text{grad } p = f(\bar{x}, t), \quad \text{div } \vec{u} = 0. \quad (6.4.1)$$

Здесь $\vec{u}$ имеет смысл вектора скорости движущейся жидкости, $p(\bar{x}, t)$ — давления, а $f(\bar{x}, t)$ — приложенной внешней силы.

Первая группа уравнений относительно компонент вектора $\vec{u}$ и давления p называется системой уравнений Навье–Стокса, а второе уравнение — уравнением неразрывности. Предполагается, что $x \in D$, где D — область из $\mathbf{R}^3$ (для простоты мы считаем ее не зависящей от t), $t \in [0, T]$ и $p(x, t)$ может быть определено с точностью до произвольной постоянной $c(t)$. Определение $\vec{u}$ и p требует кроме задания уравнений (6.4.1) также задания начальных условий, то есть $\vec{u}(x, t)$ должно быть известно при $t = 0$:

$$\vec{u}(x, 0) = \vec{u}_0(x). \quad (6.4.2)$$

При этом $\vec{u}_0(x)$ должно удовлетворять дополнительным условиям, а именно

$$\text{div } \vec{u}_0 = 0 \quad \text{и} \quad \vec{u}_0|_{\Gamma} = 0. \quad (6.4.3)$$

Кроме того, должны быть поставлены условия на границе области. Здесь существует большое разнообразие задач.

Мы упомянем лишь задание векторного поля $\vec{v}$ на границе Γ области D :

$$\vec{u}|_{\Gamma} = \vec{v}. \quad (6.4.4)$$

Одним из приемов, позволяющих получить удобную сеточную аппроксимацию, является введение дополнительного параметра ε — фиктивной вязкости. Параметр вводится так, чтобы при $\varepsilon \rightarrow \infty$ решение “параметризованной” задачи стремилось в некоторой метрике к решению исходной. Параметризованная система уравнений имеет вид [75]:

$$\frac{\partial \vec{u}_{\varepsilon}}{\partial t} - \nu \Delta \vec{u}_{\varepsilon} - \frac{1}{\varepsilon} \text{grad div } \vec{u}_{\varepsilon} + \sum_{i=1}^3 u_{i\varepsilon} \frac{\partial \vec{u}_{\varepsilon}}{\partial x_i} + \frac{1}{2} (\text{div } \vec{u}_{\varepsilon}) \vec{u}_{\varepsilon} = f. \quad (6.4.5)$$

Или в другой модификации —

$$\frac{\partial \vec{u}_\varepsilon}{\partial t} - \nu \Delta \vec{u}_\varepsilon + \sum_{i=1}^3 u_{i\varepsilon} \frac{\partial \vec{u}_\varepsilon}{\partial x_i} + \frac{1}{2} (\operatorname{div} \vec{u}_\varepsilon) \vec{u}_\varepsilon + \operatorname{grad} p_\varepsilon = f, \quad (6.4.6)$$

когда

$$\varepsilon \frac{\partial \vec{p}_\varepsilon}{\partial t} + \operatorname{div} \vec{u}_\varepsilon = 0.$$

Если в области D построена сетка с одинаковым шагом h по каждой из переменных и шагом τ по времени, то обычным способом могут быть построены разностные уравнения, имеющие порядок аппроксимации τ по времени и h^2 — по пространственным переменным:

$$U^m = F + AU^m + BU^m U^m + CU^{m-1}. \quad (6.4.7)$$

Получаем систему алгебраических уравнений с квадратичной нелинейностью.

Здесь

$$U^m = \left(u_{1(i,j,k)}^m, u_{2(i,j,k)}^m, u_{3(i,j,k)}^m \right)^T$$

m — номер момента времени, i, j, k принимают значения из множества I номеров узлов сетки, матрица

$$A = \|a_{[i,j,k;i',j',k']}^l\|, \quad (i, j, k) \in I, \quad (i', j', k') \in I; \quad l = 1, 2, 3.$$

Соответственно, B есть билинейный оператор — квадратичная форма, содержащая все произведения компонент вектора $\vec{U}$.

Приведем более подробный вид линейной составляющей уравнения (6.4.7), соответствующей первой компоненте скорости:

$$\begin{aligned} u_{1(i,j,k)}^m = & \frac{1}{\frac{1}{\tau} + \frac{6\nu}{h^2} + \frac{2}{\varepsilon h^2}} \left\{ \left(\frac{\nu}{h^2} + \frac{1}{\varepsilon h^2} \right) \left[u_{1(i+1,j,k)}^m + u_{1(i-1,j,k)}^m \right] + \right. \\ & + \frac{\nu}{h^2} \left[u_{1(i,j+1,k)}^m + u_{1(i,j-1,k)}^m + u_{1(i,j,k+1)}^m + u_{1(i,j,k-1)}^m \right] + \\ & + \frac{1}{\varepsilon h^2} \left[(u_{2(i,j+1,k)}^m - u_{2(i,j,k)}^m - u_{2(i-1,j+1,k)}^m + u_{2(i-1,j,k)}^m) + \right. \\ & + (u_{3(i,j,k+1)}^m - u_{3(i,j,k)}^m - u_{3(i-1,j,k+1)}^m + u_{3(i-1,j,k)}^m) \left. \right] + \\ & \left. + G_1^m + \frac{u_{1(i,j,k)}^{m-1}}{\tau} \right\}. \end{aligned} \quad (6.4.8)$$

G_1 включает в себя правую часть f_1 и нелинейную часть уравнения.

Уравнения, соответствующие двум другим компонентам скорости, аналогичны приведенному с точностью до циклической замены номеров компонент. Подробную запись можно найти в [49]. Слагаемое, входящее в G_l , соответствующее нелинейному члену, есть

$$\begin{aligned} \tilde{G}_{l(i,j,k)} = & \frac{h}{2(\frac{h^2}{\tau} + 6\nu + \frac{2}{\varepsilon})} \times \\ & \times \left\{ \left(u_{1(i,j,k)} + \frac{1}{2} u_{l(i,j,k)} \right) (u_{l(i+1,j,k)} - u_{l(i-1,j,k)}) + \right. \\ & + \left(u_{2(i,j,k)} + \frac{1}{2} u_{l(i,j,k)} \right) (u_{l(i,j+1,k)} - u_{l(i,j-1,k)}) + \\ & \left. + \left(u_{3(i,j,k)} + \frac{1}{2} u_{l(i,j,k)} \right) (u_{l(i,j,k+1)} - u_{l(i,j,k-1)}) \right\}. \end{aligned} \quad (6.4.9)$$

Анализ полученных выражений для коэффициентов системы разностных уравнений показывает, что за счет выбора отношения h^2/τ можно обеспечить достаточную малость суммы их модулей. Отсюда следует возможность выбора цепи Маркова, удовлетворяющей условиям согласования. Переходными плотностями могут служить абсолютные величины коэффициентов при неизвестных, входящих в систему уравнений. В предположении ограниченности нормы векторов U^m можно также за счет выбора малого шага по времени τ обеспечить стохастическую устойчивость алгоритма метода Монте-Карло нахождения U^m при заданном U^{m-1} . Здесь необходим анализ величины числа испытаний при решении системы (6.4.7) (см. гл. 5). Пробные вычисления [49] показали, что для решения трехмерных задач метод Монте-Карло в описанной выше форме может быть эффективным. Нелинейность при вычислениях учитывается путем моделирования ветвящегося процесса (см. п. 6.1). Это обеспечивает несмещенность оценок и возможность практически неограниченного параллелизма.

Недостатком является требование малости τ , в особенности при быстром изменении решения. Можно указать следующие приемы преодоления этой трудности. Надо производить дискретизацию только по пространственным переменным таким образом, чтобы искомыми были функции времени t в каждой из точек сетки. Исходную систему (6.4.7) представим в виде

$$\frac{dU}{dt} = \mathcal{K}U, \quad (6.4.10)$$

$\mathcal{K}$ — сеточный оператор с квадратичной нелинейностью. Полученную систему очень большого числа переменных можно решать известными методами: методом Эйлера, например, или более точным методом.

Если вспомнить формулы (6.4.8), (6.4.9), посредством которых определяется оператор $\mathcal{K}$, то становится ясно, что число арифметических операций на каждом шаге по времени для трехмерной сетки будет непомерно большим.

В этом случае следует определить процедуру случайного оценивания значений KU , требующую меньшего числа операций, и получать на каждом шаге по времени оценку U , несмещенную или асимптотически несмещенную (см. приложение 4). Здесь существуют различные возможности, но во всех случаях возникает весь комплекс рассмотренных нами ранее проблем: учет нелинейности, стохастическая устойчивость, возможность применения квазислучайных чисел. Подробное обсуждение всех этих проблем требует отдельной книги. Мы ограничимся лишь указанием на то, что материал, изложенный ранее, может быть инструментом их решения.

Сказанное относится и к другим подходам, основанным на дискретизации и не использующим метод искусственной вязкости.

6.5. Библиографическое дополнение к главе 6

Исследования метода Монте-Карло для решения нелинейных задач были инициированы проблемами, возникающими при расчете потоков разреженного газа. Описывающее эти потоки нелинейное уравнение содержит квадратичную нелинейность. Предложенный в [36] аналог схемы Неймана–Улама не решал проблему, поскольку мажорантный итерационный процесс в этом случае расходится, но инициировал ряд исследований [44], [61], [64]. Достаточно полный и очень хороший обзор методов решения уравнения Больцмана, основанных на прямом моделировании движения частиц (молекул), содержит монография Рязанова и Вагнера [25].

Методы, основанные на моделировании ветвящихся процессов, применялись при решении нелинейных задач теплопроводности [46].

Возможность использования метода Монте-Карло для линейных систем, возникающих в схеме последовательных приближений для решения нелинейных уравнений, неоднократно отмечалась в литературе (в частности, в [6], [40]). В недавней работе Холтона [14] приведен ряд примеров решения нелинейных уравнений рандомизованным методом Ньютона (без исследования вопросов стохастической устойчивости).

В последнее время было обнаружено (см. п. 6.3 и П.4), что оптимальные в детерминированном смысле методы после рандомизации могут приводить к процедурам, не самым лучшим в смысле устойчивости и трудоемкости.

Заключение

Читатель, который хотя бы бегло просмотрел эту книгу, несомненно заметит, что в конечном счете речь идет о вычислении интегралов по траекториям марковских цепей. Интегралы вычисляются методом Монте-Карло (в самом общем случае) или методом *квази*Монте-Карло (при некоторых ограничениях). Ясно, что коль скоро искомое решение представлено в виде интеграла, то его аппроксимацией является сумма, алгоритм вычисления которой обладает высокой степенью параллелизма. Как мы видим, решение линейных систем с большим числом неизвестных полезно искать с помощью квадратурных (кубатурных) формул! Высказанная точка зрения намечает достаточно универсальный подход к конструированию параллельных алгоритмов и эффективному использованию компьютеров с очень большим числом процессоров.

В книге, по-видимому, впервые с точки зрения параллелизма исследуется случай, когда решение системы не представимо в виде интеграла ($\lambda_1(|A|) > 1$), но представимо в виде повторного интеграла. Указан способ эффективного использования квазислучайных чисел в этом случае.

Безусловно, книга отражает лишь очень небольшую область применения методов Монте-Карло и *квази*Монте-Карло, но это позволило сосредоточиться на некоторых принципиально важных вопросах и провести параллели между детерминированными и стохастическими методами.

Можно отметить, что материал книги подтверждает точку зрения, согласно которой методы математической статистики позволяют сделать выводы о свойствах очень большой совокупности по выборочным данным. Не нужно решать итерационными методами системы с 10^9 и более неизвестными, если результат нужно получить с двумя значащими цифрами. С другой стороны, не следует использовать статистические методы при решении систем небольшой размерности или при вычислении однократного интеграла от гладкой функции.

Ближкие соображения несомненно имеют место и при решении широкого круга интегральных уравнений.

В заключение следует сказать, что со спецификой численного решения стохастических дифференциальных уравнений, краевых задач для уравнений в частных производных и многими приложениями метода Монте-Карло читателю придется знакомиться по другим монографиям, благо их немало. И нам остается пожелать заинтересованному читателю приятного чтения.

Приложения

П.1. Простейшие оценки и условия согласования *)

Прямые оценки

$$\tau = 0, 1, \dots;$$

$$k_{-1,0} = 1, h_0 = h(x_0),$$

$$k_{j-1,j} = k(x_{j-1}, x_j); f_j = f(x_j).$$

$$\text{Соответственно } p_0^0 = p^0(x_0)$$

$$J_a = \frac{h_0 k_{0,1} \dots k_{\tau-1,\tau}}{p_0^0 p_{0,1} \dots p_{\tau-1,\tau}} \cdot \frac{f_\tau}{g_\tau}$$

$$J_c = \sum_{l=0}^{\tau} \frac{h_0 k_{0,1} \dots k_{l-1,l} f_l}{p_0^0 p_{0,1} \dots p_{l-1,l}}$$

Условия согласования для прямых оценок

Для J_a :

$$p^0(x) > 0, \text{ если } h(x) \neq 0$$

$$p(x, y) > 0, \text{ если } k(x, y) \neq 0$$

$$g(x) > 0, \text{ если } f(x) \neq 0$$

Для J_c :

$$p^0(x) > 0, \text{ если } h(x) \neq 0$$

$$p(x, y) > 0, \text{ если } k(x, y) \neq 0$$

Сопряженные оценки

$$J_a^* = \frac{f_0 k_{1,0} \dots k_{\tau,\tau-1}}{p_0^0 p_{0,1} \dots p_{\tau-1,\tau}} \cdot \frac{h_\tau}{g_\tau}$$

$$J_c^* = \sum_{l=0}^{\tau} \frac{f_0 k_{1,0} \dots k_{l-1,l} f_l}{p_0^0 p_{0,1} \dots p_{l-1,l}}$$

Условия согласования для сопряженных оценок

Для J_a :

$$p^0(x) > 0, \text{ если } f(x) \neq 0$$

$$p(x, y) > 0, \text{ если } k(y, x) \neq 0$$

$$g(x) > 0, \text{ если } h(x) \neq 0$$

Для J_c :

$$p^0(x) > 0, \text{ если } f(x) \neq 0$$

$$p(x, y) > 0, \text{ если } k(y, x) \neq 0$$

*) Траектории процесса обозначены как $x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_\tau$, $x_j \in X$; и в случае линейных алгебраических уравнений $i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_\tau$, $i_j = 1, \dots, n$; $j = 0, \dots, \tau$.

П.2. Таблица простых чисел

[illegible]

П.3. Таблица числителей $r_j^{(l)}$, [71]

$j \backslash l$	1	2	3	4	5	6	7	8	9	10	11
1	1	1	1	1	1	1	1	1	1	1	1
2	1	3	5	15	17	51	85	255	257	771	1285
3	1	1	7	11	13	61	67	79	465	721	823
4	1	3	7	5	7	43	49	147	439	1013	727
5	1	1	5	3	15	51	125	141	177	759	267
6	1	3	1	1	9	59	25	89	321	835	833
7	1	1	3	7	31	47	109	173	181	949	471
8	1	3	3	9	9	57	43	43	225	113	1601
9	1	3	7	13	3	35	89	9	235	929	1341
10	1	1	5	11	27	53	69	25	103	615	913
11	1	3	5	1	15	19	113	115	411	157	1725
12	1	1	7	3	29	51	47	97	233	39	2021
13	1	3	7	7	21	61	55	19	59	761	1905
14	1	1	1	9	23	37	97	97	353	169	375
15	1	3	3	5	19	33	3	197	329	983	893
16	1	1	3	13	11	7	37	101	463	657	1599
17	1	1	7	13	25	5	83	255	385	647	415
18	1	3	5	11	7	11	103	29	111	581	605
19	1	1	1	3	13	39	27	203	475	505	819
20	1	3	1	15	17	63	13	65	451	833	975
21	1	1	5	5	1	27	33	195	263	139	915
22	1	3	3	3	25	17	115	177	19	147	1715
23	1	1	3	15	29	15	41	105	249	203	1223
24	1	3	1	7	3	23	79	17	275	81	1367
25	1	3	7	9	31	29	17	47	369	337	663

В таблице приведены числители $r_j^{(l)}$ направляющих чисел при $1 \leq j \leq 51$, $1 \leq l \leq 20$. Направляющие числа $v_j^{(l)}$ рассчитываются по формуле. $v_j^{(l)} = r_j^{(l)} 2^{-l}$, последовательность векторов $Q_n = (q_{n,1}, \dots, q_{n,s})$ рассчитывается по формулам на с. 94.

12	13	14	15	16	17	18	19	20
1	1	1	1	1	1	1	1	1
3855	4369	13107	21845	65535	65537	196611	327685	983055
4091	4125	4141	28723	45311	53505	250113	276231	326411
987	5889	6915	16647	49925	116487	83243	116529	715667
1839	6929	16241	16565	17139	82207	50979	252717	851901
4033	3913	11643	18777	35225	102401	45059	36865	299009
2515	6211	2147	3169	35873	33841	99889	247315	1032727
579	1731	11977	7241	63609	81003	15595	144417	685617
3863	1347	4417	5087	12631	103445	152645	130127	775365
977	6197	14651	2507	27109	5205	91369	302231	172023
3463	2817	9997	7451	12055	44877	24895	508255	574033
2909	5459	2615	13329	35887	97323	83101	320901	810643
3379	8119	13207	8965	9997	75591	226659	187499	628265
1349	5121	13313	19457	1033	62487	250917	234593	308321
3739	7669	2671	18391	31161	12111	259781	36159	232401
347	2481	5201	3123	32253	78043	63447	508757	974837
387	7101	11469	11699	15865	49173	147489	81991	802875
2381	2677	14855	721	26903	100419	206167	241771	987201
2821	1405	12165	709	41543	57545	77163	357231	378135
1873	7423	5837	20481	12291	86017	12303	299025	774207
1959	725	5387	19285	5165	27985	69809	128325	164575
1929	2465	12483	13057	28931	54019	21251	62233	248081
2389	471	12945	32321	29377	127427	103759	472541	1008719
3251	2887	1279	4865	64771	24321	42247	338691	599831
1149	1715	187	12285	53631	110851	4357	153347	671033

$j \backslash l$	1	2	3	4	5	6	7	8	9	10	11
26	1	1	5	13	11	3	29	169	393	829	629
27	1	3	1	9	5	21	119	109	167	989	525
28	1	1	3	1	23	13	75	149	333	375	469
29	1	3	3	11	27	31	73	15	473	365	981
30	1	1	7	7	19	25	105	213	469	131	1667
31	1	3	5	5	21	9	7	135	101	215	1587
32	1	1	1	15	5	49	59	253	21	733	1251
33	1	1	1	1	1	33	65	191	451	451	451
34	1	3	5	15	17	19	21	155	229	447	481
35	1	1	7	11	13	29	3	175	247	177	721
36	1	3	7	5	7	11	113	63	297	57	483
37	1	1	5	3	15	19	61	47	403	471	1209
38	1	3	1	1	9	27	89	7	497	979	1457
39	1	1	3	7	31	15	45	23	61	197	415
40	1	3	3	9	9	25	107	39	361	251	1435
41	1	3	7	13	3	3	25	55	215	517	725
42	1	1	5	11	27	21	5	71	393	137	861
43	1	3	5	1	15	51	49	87	125	567	41
44	1	1	7	3	29	19	111	103	285	1021	1619
45	1	3	7	7	21	29	119	119	501	167	1579
46	1	1	1	9	23	5	33	135	277	877	1701
47	1	3	3	5	19	1	67	153	199	929	869
48	1	1	3	13	11	39	101	169	301	269	1151
49	1	1	7	13	25	37	19	185	19	327	1897
50	1	3	5	11	7	43	39	201	83	997	1679
51	1	1	1	3	13	7	91	217	351	91	1355

12	13	14	15	16	17	18	19	20
243	5595	8133	4929	10817	8261	189901	255947	734787
3609	5689	11819	15889	48083	67537	63993	336469	749285
1131	441	14471	12625	8881	34707	85105	479495	911133
1701	3169	7615	8405	41135	106823	107847	339031	977907
143	4485	2981	12593	60913	1573	26967	507907	344073
1339	6311	4081	28637	60935	94129	109273	475921	281389
3497	3557	7223	13425	58577	69521	217151	424277	789985
2499	483	11843	28285	12029	86021	217093	348165	176165
1571	3781	10799	15893	959	19793	213491	377941	414943
983	3195	9277	15405	19637	87283	186143	343297	1041185
4021	5213	2031	4677	26607	20931	54345	259163	741087
1625	5085	15371	19493	56445	26369	27399	521499	132383
3217	185	6603	1129	36087	66817	98051	451841	175361
1163	7323	7563	25321	52563	37745	81777	235347	539895
2977	1713	11617	14979	5455	68289	209987	346179	521289
3391	4021	4129	4099	12345	102733	21287	128115	20689
675	5875	12061	25469	47423	29505	124097	444613	430923
3093	5363	3471	17589	50131	33137	98739	361365	426737
1495	4977	15919	6731	43771	23313	151281	270519	11187
3443	5441	1097	13483	58779	36561	116819	420599	998391
557	1779	10369	15325	33331	118321	59665	498897	494137
675	6777	14343	18465	63615	43349	30799	322567	939017
1489	287	8475	6929	46013	52785	75249	14035	507165
2303	6919	16139	16677	34579	120981	239693	73299	863545
3925	1517	305	21765	45827	91157	113679	204881	761911
3705	1875	7621	4381	9079	94533	37261	431301	176455

П.4. Примеры вычислений ([48], [76])

В конце главы 6 упоминались три метода построения оценок решения системы алгебраических уравнений с квадратичной нелинейностью:

- (1) стохастический метод Ньютона;
- (2) стохастический метод, основанный на простой линеаризации;
- (3) стохастический вариант метода искусственного хаоса.

Все эти методы являются итерационными, причем на каждой итерации (будем называть их испытаниями) необходимо строить оценку функционалов от решения некоторой системы линейных уравнений. Эта оценка строилась с помощью сопряженной оценки по столкновениям J_c^* . Причем в случае решения системы уравнений вида

$$x_i = \sum_{j=1}^d a_{i,j} x_j + f_i, \quad i = 1, \dots, d, \quad (*)$$

выбирались следующие параметры марковской цепи:

- начальное распределение задается вероятностями выбора начальной точки i : $\pi_i = |f_i| / \sum_{i=1}^d |f_i|$;
- вероятность поглощения не зависит от состояния марковской цепи: $g = 0.05$;
- вероятность перехода из состояния j в состояние i : $p_{j \rightarrow i} = \frac{1}{1-g} |a_{i,j}| / \sum_{i=1}^d |a_{i,j}|$.

Напомним, что сопряженная оценка по столкновениям задается следующим образом. Пусть x — решение системы (*); h — произвольный вектор из $\mathbf{R}^d$; $\xi_0, \xi_1, \xi_2, \dots$ — марковская цепь с заданными выше характеристиками; $\tau+1$ — момент поглощения марковской цепи. Тогда в случае выполнения некоторых условий случайная величина

$$J_c^* = \sum_{k=0}^{\tau} \frac{f_{\xi_0}}{\pi_{\xi_0}} \frac{a_{\xi_1, \xi_0}}{p_{\xi_0 \rightarrow \xi_1}} \dots \frac{a_{\xi_k, \xi_{k-1}}}{p_{\xi_{k-1} \rightarrow \xi_k}} h_{\xi_k}$$

является несмещенной оценкой скалярного произведения векторов h и x .

В качестве результата численного эксперимента можно привести графики логарифмов оценок математического ожидания нормы невязки совместно с границами доверительных интервалов для этих оценок (см. ниже примеры 1–3). Для построения оценки математического ожидания в каждом случае моделировалось по 100 реализаций оценок. Границы доверительных интервалов в нашем случае “сливаются” на графиках с самой оценкой (рис. П.1–П.3).

Пример 1 — стохастический метод Ньютона лучшеРешается система уравнений ^{*)}

$$x_1 = 0.3013 + 0.0881x_1 + 0.0846x_2 + 0.0949x_1^2 + 0.6012x_1x_2 + 0.2677x_2^2,$$

$$x_2 = 0.0259 + 0.0247x_1 + 0.0703x_2 + 0.0258x_1^2 + 0.1492x_1x_2 + 0.1114x_2^2.$$

Интересующее нас решение есть $(x_1, x_2) = (0.3588, 0.0437)$.

Для построения оценок функционалов от решений вспомогательных линейных систем на каждом испытании каждого метода моделировалось по 100 реализаций марковских цепей. Видно, что в данном примере лидером показал себя стохастический метод Ньютона. Горизонтальный участок на графике соответствует периоду, когда для точных вычислений не хватало разрядности чисел.

Логарифм оценки математического ожидания нормы невязки

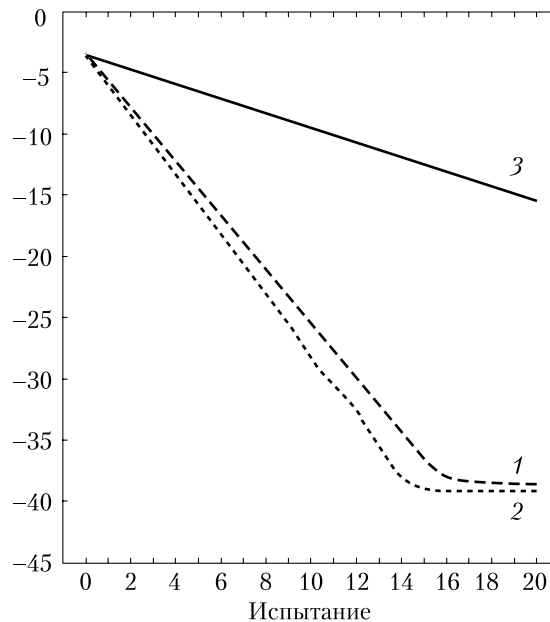


Рис. П.1

1 — простой метод, 2 — метод Ньютона,
3 — метод искусственного хаоса

^{*)} Коэффициенты уравнений и компоненты их решений приведены с точностью 10^{-4} .

Пример 2 — простая линеаризация лучше

Решается система уравнений

$$x_1 = 0.3013 + 0.3525x_1 + 0.3386x_2 + 0.0237x_1^2 + 0.1504x_1x_2 + 0.0669x_2^2$$

$$x_2 = 0.0259 + 0.0989x_1 + 0.2811x_2 + 0.0065x_1^2 + 0.0374x_1x_2 + 0.0278x_2^2.$$

Интересующее нас решение есть $(x_1, x_2) = (0.5559, 0.1192)$.

Для построения оценок функционалов от решений вспомогательных линейных систем на каждом испытании каждого метода моделировалось по 5 реализаций марковских цепей. В данном примере лидером показал себя стохастический метод, основанный на простой линеаризации.

Логарифм оценки математического ожидания нормы невязки

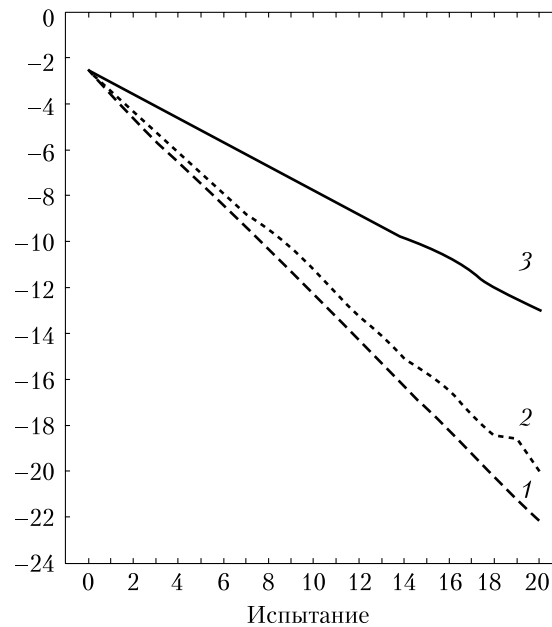


Рис. П.2

1 — простой метод, 2 — метод Ньютона,
3 — метод искусственного хаоса

Пример 3 — метод искусственного хаоса лучше

Решается система уравнений

$$x_1 = 0.3013 + 0.2203x_1 + 0.2116x_2 + 0.0012x_1^2 + 0.0076x_1x_2 + 0.0033x_2^2,$$

$$x_2 = 0.0259 + 0.0618x_1 + 0.1757x_2 + 0.0003x_1^2 + 0.0018x_1x_2 + 0.0014x_2^2.$$

Интересующее нас решение есть $(x_1, x_2) = (0.4038, 0.0618)$.

Для построения оценок функционалов от решений вспомогательных линейных систем на каждом испытании каждого метода моделировалась только одна реализация марковской цепей. Здесь бóльшую скорость сходимости демонстрирует метод искусственного хаоса.

Логарифм оценки математического ожидания нормы невязки

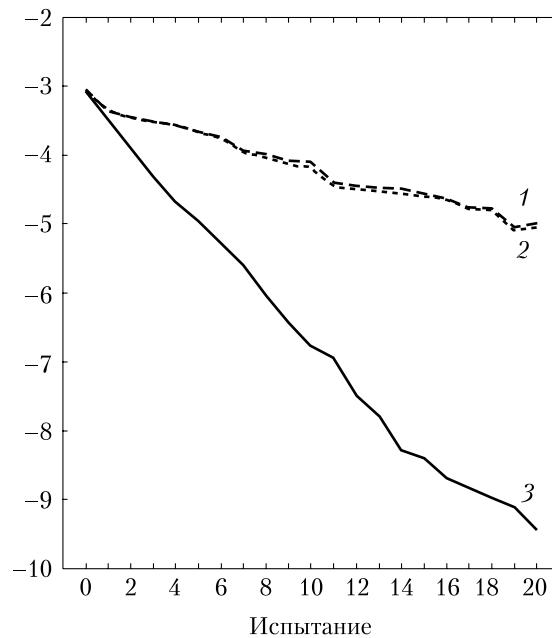


Рис. П.3

1 — простой метод, 2 — метод Ньютона,
3 — метод искусственного хаоса

Ниже описан вычислительный эксперимент решения с помощью рандомизации оператора $\mathcal{K}$ системы (6.4.10). Система получена как разностный аналог уравнений Навье–Стокса с введенной искусственной вязкостью, описывающей движение вязкой несжимаемой жидкости:

$$\frac{\partial v}{\partial t} + (v \nabla) v = \frac{1}{\rho \varepsilon} \operatorname{grad} \operatorname{div} v - \frac{1}{2} v \operatorname{div} v + \nu \Delta v.$$

Здесь $v = v(t, x) = (v_1(t, x), v_2(t, x), v_3(t, x))$ — векторное поле скоростей в трехмерном пространстве ($x = (x_1, x_2, x_3) \in \mathbf{R}^3$). Величины ρ и ν являются скалярными параметрами, определяющими плотность и вязкость жидкости соответственно. Коэффициент ε — параметр искусственной вязкости. Оператор $\mathcal{K}$ определяется уравнениями (6.4.8), (6.4.9). Для учета нелинейности использовалась поправка, получаемая в результате разложения в ряд Тейлора квадратичной функции случайных переменных в окрестности их математических ожиданий.

Приведенное уравнение использовалось для расчета движения жидкости в кубической камере с подвижной боковой стенкой:

$$x \in D, \text{ где } D = [0, 1]^3.$$

В начальный момент времени ($t = 0$) жидкость полагалась неподвижной:

$$v(0, x) = 0.$$

На границах задавалось условие прилипания, а именно:

$$v|_{\partial D}(x, t) = \begin{cases} (0, 1 - \frac{1}{1-t^5}, 0), & x_1 = 0; \\ (0, 0, 0), & x_1 \neq 0. \end{cases}$$

Пространственная сетка содержала $l \times l \times l$ точек — одинаковое число по каждой переменной. Использовалась стандартная аппроксимация производных с точностью до h^2 , $h = 1/l$. Полагалось $\varepsilon = 0.5$, $\rho = 1$, ν в разных примерах меняется от 0.0001 до 0.25.

Вычислялись оценки решения в момент времени $t = 1$ и полагалось $\Delta t = 1/N$. Обычные детерминированные вычисления по явной схеме налагают соответствующие условия на соотношения h и Δt вида $\Delta t \leq ch^2$, $c = \text{const}$ обеспечения устойчивости относительно ошибок округления. Во всех случаях эти условия были выполнены. Кроме того, рандомизация требует выполнения условий стохастической устойчивости. Они уменьшают величину константы c .

В случае использования детерминированного метода Эйлера необходимо провести $3(l-2)^3 L$ расчетов отдельных компонент функции v . Если использовать при расчетах рандомизацию, состоящую в том, что на каждом шаге вычислялись значения v в случайно выбранной точке сетки, то

для фиксированного числа усреднений m оценки $\hat{F}$ на каждом шаге получим, что для расчета оценки решения к моменту времени 1 потребуется mL расчетов компонент $\mathcal{K}$. При этом мы предполагаем, что величина m должна быть достаточной для обеспечения стохастической устойчивости и, следовательно, распараллеливания алгоритма. Как показывают численные исследования, в данном примере наименьшее значение параметра m , при котором сохраняется стохастическая устойчивость, убывает с уменьшением Δt (вплоть до 1!). Это сближает описанную методику с методами, использующими аппарат стохастических дифференциальных уравнений. Таким образом, существует t_0 , и в случае $0 < \Delta t < t_0$ рандомизованный метод Эйлера быстрее детерминистического в $t_0/\Delta t$ раз, если не учитывать случайную погрешность.

В следующей таблице для ряда параметров l и Δt приведены наименьшие значения m , при которых сохраняется стохастическая устойчивость, а также показано, во сколько раз сокращается расчет компонент вектора $\hat{F}$ при использовании рандомизованного метода Эйлера по сравнению детерминистическим.

l	Δt	Наименьшее m	$3(l-2)^3/m$ – преимущество рандомизованного метода
10	0.001	461	3.33
10	0.0005	183	8.39
10	0.0001	19	80.80
10	0.00005	1	1536.00
20	0.0001	1731	10.12

Видно, что скорость расчета оценок при помощи рандомизованного метода Эйлера может оказаться существенно выше, чем скорость расчета при помощи детерминистического метода Эйлера. Однако при расчетах рандомизованным методом Эйлера возникает дополнительная случайная погрешность. Влияние случайной погрешности может быть снижено за счет усреднения нескольких реализаций оценки.

Интерес представляет также поведение погрешности оценки, получаемой путем усреднения нескольких независимых оценок решения в момент времени t , расчет которых мог бы проводиться на разных компьютерах. Эксперименты показывают, что для достаточно большого диапазона числа усредняемых оценок N (при расчетах N бралось от 1 до 512) норма отклонения усредненной оценки решения от оценки, получаемой детерминистическим методом Эйлера, убывает как $c/\sqrt{N}$, где $c = \text{const}$.

Можно сделать вывод о целесообразности параллельного проведения расчетов на нескольких независимых вычислителях с дальнейшим усреднением результатов.

Литература

- [1] Bahr B., Esseen C-G. Inequalities for the r -th absolute moment of a sum of random variables ($1 \leq r \leq 2$) // Ann. of the Math. Stat. 1965. Vol. 36. № 1. P. 229–303.
- [2] Bird G. A. Molecular Gas Dynamics. London: Oxford University Press, 1976. 458 p.
- [3] Coveyou R. R., Makpherson R. P. Fourier analysis of uniform random number generators // J. of Assoc. for Comp. Math. 1967. Vol. 14. № 1. P. 100–119.
- [4] De Vroye L. Non-Uniform random variate generation. Berlin, 1986. 624 p.
- [5] Ermakov S. M. Stochastical stability, Neumann–Ulam scheme and particle methods // IVth IMACS Seminar on Monte Carlo Methods: Proceedings (September 15–19, 2003, Berlin). Berlin: WIAS. 2003. P. 55.
- [6] Ermakov S., Kaloshin I. Solving the Nonlinear Algebraic Equations with Monte Carlo Method // Advances in Stochastic Simulation Methods. Boston–Basel–Berlin: Birkhauser, 2000. P. 3–15.
- [7] Ermakov S. M., Wagner W. Monte Carlo difference schemes for the wave equations // Monte Carlo Methods and Appl. 2002. Vol. 8. № 1. P. 1–29.
- [8] Fishman G. S. Monte Carlo Concept, Algorithms and applications. Springer, 1996. 698 p.
- [9] Fishman G. S. Principles of discrete Event Simulation. N.Y.: John Wiley & Sons, 1978. XVII+514 p.
- [10] Franklin J. N. Deterministic simulation of random processes // Math. Comp. 1963. Vol. 17. P. 28–59.
- [11] Gerlovina V., Nekrutkin V. Asymptotical behavior of linear congruential generators // Monte Carlo Methods and Appl. 2005. Vol. 11. № 2. P. 135–162.
- [12] Glasserman P. Monte Carlo Methods in Financial Engeneering. Springer-Verlag, 2003. 596 p.
- [13] Halton J. H. Sequential Monte Carlo // Proc. Cambr. Phil. Soc. 1962. Vol. 58. № 1. P. 57–58.
- [14] Halton J. H. Sequential Monte Carlo techniques for solving non-linear systems // Monte Carlo Methods And Applications. 2006. Vol. 12. № 2. P. 113–141.
- [15] Hammersley J. M., Handscomb D. C. Monte Carlo methods. New York – London – Methuen: Jonh Wiley & Sons, 1964. 198 p.
- [16] Hsu L. C. A general approximation method of evaluating multiple integrals // Tohoi math. J. 1957. Vol. 9. №1. P. 45–55.

- [17] Jansson B. Random number generators. Stockholm: Victor Pettersons Bokindustriab, 1966. 212 p.
- [18] Кас М. On some connections between probability theory and differential and integral equations // Proceedings of the Second Berkeley Symposium on Probability Theory and Mathematical Statistics (1950). Berkeley: University of California Press, 1951. P. 189–215.
- [19] Kalos M. H., Whitlock P. A. Monte Carlo methods. N.Y.: John Wiley & Sons, 1986. 208 p.
- [20] Knuth D. E., Yao A. C. The Complexity of Nonuniform Random Number Generation // Algorithms and Complexity. N. Y.: Academic press, 1976. P. 357–428.
- [21] Lehmer P. H. Mathematical methods in large-scale computing units // Ann. Comp. Lab., Harvard university. 1951. Vol. 26. P. 141–146.
- [22] Liu J. S. Monte Carlo strategies in scientific Computing. Springer, 2003. 359 p.
- [23] Marsaglia G. Random numbers fall mainly in the planes // U.S.A.: Proc. National Academy of Sciences, 1968. Vol. 61. P. 25–28.
- [24] Niederreiter H. Random Number Generation and Quasi-Monte Carlo Methods. CBMS-NSF Regional Conference Series in Applied Mathematics, 63. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. 244 p.
- [25] Rjasanow S., Wagner W. Stochastic Numerics for the Boltzmann Equation. Berlin; Heidelberg: Springer-Verlag, 2005. 256 p.
- [26] Tauswothe R. C. Random numbers generated by linear recurrence modul two // Math. Comp. 1965. Vol. 19. № 90. P. 201–209.
- [27] Walker A. J. New Fast method for generating discrete random numbers with arbitrary friguency distributions // Elektronik Letters. 1974. Vol. 10. P. 127–128.
- [28] Бахвалов Н. С. Об оптимальных оценках сходимости квадратных процессов и методов интегрирования типа Монте-Карло на классах функций // Численные методы решения дифференциальных и интегральных уравнений: Сб. М.: Наука, 1964. С. 5–63.
- [29] Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. Численные методы. М.: Физматлит; СПб.: Невский Диалект, 2001. 630 с.
- [30] Вагнер В., Ермаков С. М. Стохастическая устойчивость и параллелизм метода Монте-Карло // ДАН. 2001. Т. 379. № 4. С. 439–441.
- [31] Гихман И. И., Скороход Ф. В. Введение в теорию случайных процессов. М.: Наука, 1977. 568 с.

- [32] Гладкова Л. А., Ермаков С. М. Рекуррентные алгоритмы Монте-Карло для решения кинетических уравнений // Статистические модели с приложениями в эконометрике и смежных областях: Сб. СПб.: Изд-во НИИХ СПбГУ, 1999. С. 50–75.
- [33] Голуб Дж., Ван Лоун У. Матричные вычисления. М.: Мир, 1999. 548 с.
- [34] Егоров А. Д., Соболевский П. И., Янович Л. А. Приближенные методы вычисления континуальных интегралов. Минск: Наука и Техника, 1985. 310 с.
- [35] Ермаков С. М. Метод Монте-Карло и смежные вопросы. М.: Наука, 1975. 472 с.
- [36] Ермаков С. М. Об аналоге схемы Неймана–Улама в нелинейном случае // ЖВМ и МФ. 1973. Вып. 13. № 3. С. 564–573.
- [37] Ермаков С. М., Беляева А. А. О методе Монте-Карло с запоминанием промежуточных результатов // Вестник СПбГУ. Сер. Математика, механика, астрономия. 1996. Вып. 29. № 3. С. 5–8.
- [38] Ермаков С. М., Иванова А. В. О стохастической сutoйчивости разностных схем // Вестник СПбГУ. Сер. Математика, механика, астрономия. 1991. Вып. 1. № 1. С. 30–34.
- [39] Ермаков С. М., Калошин И. В., Тимофеев К. А. Метод искусственного хаоса для решения методом Монте-Карло уравнений с квадратичной нелинейностью // Математические модели. Теория и приложения: Сб. Вып. 7. СПб.: Изд-во НИИХ СПбГУ, 2006. С. 3–20.
- [40] Ермаков С. М., Ликинова О. М., Калошин И. В. Об одной итерационной схеме решения задач с квадратичной нелинейностью // Математические модели. Теория и приложения: Сб. Вып. 6. СПб.: Изд-во НИИММ, 2005. С. 26–33.
- [41] Ермаков С. М., Мелас В. Б. Математический эксперимент с моделями сложных стохастических систем. СПб.: Изд-во СПбГУ, 1993. 270 с.
- [42] Ермаков С. М., Мисов Т. И. Моделирование Δ^2 -распределения // Вестник СПбГУ. Сер. 1. 2005. Вып. 4. С. 53–60.
- [43] Ермаков С. М., Михайлов Г. А. Статистическое моделирование. М.: Наука, 1982. 296 с.
- [44] Ермаков С. М., Москалева Н. М. Ветвящиеся процессы и уравнение Больцмана. Вычислительные аспекты // Вестник ЛГУ. Сер. 1. 1987. Вып. 3(15). С. 38–43.
- [45] Ермаков С. М., Некруткин В. В., Сипин А. С. Случайные процессы для решения классических задач математической физики. М.: Наука, 1984. 206 с.
- [46] Ермаков С. М., Расулов А. С., Бакаев М. Т., Веселовская А. З. Избранные алгоритмы метода Монте-Карло. Ташкент: Университет, 1992. 132 с.

- [47] Ермаков С. М., Рукавишникова А. И. Квази Монте-Карло алгоритмы решения систем линейных алгебраических уравнений // Математические модели. Теория и приложения: Сб. Вып. 6. СПб.: Изд-во ВВМ, 2005. С. 3–26.
- [48] Ермаков С. М., Рукавишникова А. И., Тимофеев К. А. О некоторых стохастических и квазистохастических методах решения уравнений // Вестник СПбГУ. Сер. 1. 2008. Вып. 4. С. 43–51.
- [49] Ермаков С. М., Шакенов К. К. О применении метода Монте-Карло к разностному аналогу уравнений Навье–Стокса // Вестник ЛГУ. 1986. 13 с. (ВИНИТИ, номер регистрации 8576-8866.)
- [50] Золотухин В. Г., Ермаков С. М. Применение метода Монте-Карло к расчету защиты от ядерных излучений. Вопросы физической защиты реакторов. М.: Госатомиздат, 1963. С. 171–182.
- [51] Калашин И. В. <http://vega.math.spbu.ru/igor/stability/>
- [52] Кейперс Л., Нидеррейтер Г. Равномерное распределение последовательностей. М.: Наука, 1985. 408 с.
- [53] Кнут Д. Искусство программирования для ЭВМ. Т. 2: Получисленные алгоритмы. М.: Мир, 1997. 724 с.
- [54] Кнут Д., Яо Э. Сложность моделирования неравномерных распределений. М.: Мир, 1983. С. 97–158. (Кибернетический сборник. Новая серия. Вып. 19.)
- [55] Коробов Н. М. Теоретико-числовые методы в приближенном анализе. М.: Физматгиз, 1963. 224 с.
- [56] Марчук Г. И. и др. Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, 1976. 283 с.
- [57] Михайлов Г. А. Весовые методы Монте-Карло. Новосибирск, 2000, 248 с.
- [58] Михайлов Г. А. Оптимизация весовых методов Монте-Карло. М.: Наука, 1987. 240 с.
- [59] Михайлов Г. А., Войтишек А. В. Численное статистическое моделирование. Методы Монте-Карло. М.: Академия, 2006. 367 с.
- [60] Некруткин В. В. К обоснованию метода отбора для мультипликативного датчика псевдослучайных чисел // ЖВМ и МФ. 1981. Вып. 2. № 5. С. 1100–1107.
- [61] Некруткин В. В. Прямая и сопряженная схема Неймана–Улама для решения нелинейных интегральных уравнений // ЖВМ и МФ. 1974. Вып. 14. № 6. С. 1409–1415.
- [62] Ортега Д., Рейнболдт В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир, 1975. 558 с.

- [63] Пригарин С. М. Методы численного моделирования случайных процессов и полей. Новосибирск: Изд-во ИВМ и МГ СО РАН, 2005. 259 с.
- [64] Расулов А. С. Метод Монте-Карло для решения нелинейных задач. Ташкент: ФАН, 1992. 104 с.
- [65] Сабельфельд К. К. Методы Монте-Карло в краевых задачах. Новосибирск: Наука, 1989. 280 с.
- [66] Самарский А. А., Вабищевич П. Н., Матус П. П. Разностные схемы с операторными множителями. Минск, 1998. 442 с.
- [67] Севастьянов Б. А. Ветвящиеся процессы. М.: Наука, 1971. 436 с.
- [68] Синай Я. Г. Введение в эргодическую теорию. М.: Фазис, 1996. 128 с.
- [69] Соболев С. Л. Введение в теорию кубатурных формул. М.: Наука, 1974. 808 с.
- [70] Соболев И. М. Многомерные квадратурные формулы и функции Хара. М.: Наука, 1969. 288 с.
- [71] Соболев И. М., Статников Р.Б. Выбор оптимальных параметров в задачах со многими критериями. М.: Дрофа, 2006. 226 с.
- [72] Спанье Дж., Э. Гелбард. Метод Монте-Карло и задачи переноса нейтронов. М.: Атомиздат, 1972. 271 с.
- [73] Стронгин Р. Г. Численные методы в многоэкстремальных задачах. М.: Наука, 1978. 240 с.
- [74] Таусворт. Случайные числа, порождаемые линейными рекуррентными соотношениями по модулю 2. М.: Мир, 1979. С. 62–73. (Кибернетический сборник. Новая серия. Вып. 16.)
- [75] Темам Р. Уравнения Навье–Стокса. Теория и численный анализ. М.: Мир, 1981. 408 с.
- [76] Тимофеев К. А. Об одном классе методов Монте-Карло для решения уравнений с квадратичной нелинейностью уравнений // Вестник СПбГУ. Сер. 10. Вып. 3. 2008. С. 23–28.
- [77] Уилкс С. Математическая статистика. М.: Наука, 1967. 632 с.
- [78] Успенский В. А., Семенов А. Л. Теория алгоритмов: основные открытия и приложения. М.: Наука, 1985. 408 с.
- [79] Фролов А. С., Ченцов Н. Н. Решение трех типичных задач теории переноса методом Монте-Карло // Метод Монте-Карло в проблеме переноса излучений: Сб. М.: Атомиздат, 1967. С. 25–52.
- [80] Фролов А. С., Ченцов Н. Н. О вычислении методом Монте-Карло определенных интегралов, зависящих от параметра // ЖВМ и МФ. 1962. Вып. 2. № 4. С. 714–717.
- [81] Хисамутдинов А. И. Единичный класс оценок для вычисления по методу Монте-Карло функционалов от решения интегрального уравнения 2-го рода // ЖВМ и МФ. 1970. Вып. 10. № 5. С. 1269–1280.