

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
КАФЕДРА ТЕОРИИ СИСТЕМ УПРАВЛЕНИЯ ЭЛЕКТРОФИЗИЧЕСКОЙ АППАРАТУРОЙ

**Сетина Ирина Андреевна**

**Магистерская диссертация**

**Оптимальное планирование эксперимента  
для экспоненциальной регрессии**

Направление 03.04.01

Прикладные математика и физика

Магистерская программа математические и информационные технологии

Научный руководитель,  
кандидат физ.-мат. наук,  
доцент  
Владимирова Л. В.

Санкт-Петербург

201

## Содержание

Введение .....	3
Постановка задачи .....	5
Глава I. Обзор литературы .....	6
Глава II. Оптимальный план для регрессионного эксперимента .....	10
2.1. Классическая схема регрессии .....	10
2.2. Нелинейная экспоненциальная регрессия .....	11
2.3. Локально $D$ -оптимальные планы для экспоненциальной регрессии .....	12
Глава III. Метод случайного поиска глобального экстремума на основе оценивания элементов матрицы ковариаций .....	14
3.1. Простейший случайный поиск (слепой поиск) .....	14
3.2. Доказательство сходимости простейшего случайного поиска .....	15
3.3. Оценивание числа испытаний .....	16
3.4. Случайный поиск с “памятью” .....	16
3.5. Введение неравномерного распределения .....	17
3.6. Алгоритм метода случайного поиска с “памятью” .....	18
Глава IV. Метод Метрополиса для нахождения моды плотности распределения случайных величин .....	25
4.1. Выборка с отклонением .....	25
4.2. Модифицированная выборка с отклонением .....	26
4.3. Метод Метрополиса .....	27
4.4. Частный случай алгоритма Метрополиса .....	28
Глава V. Численные результаты .....	30
Выводы .....	33
Заключение .....	34
Список литературы .....	35
Приложения .....	39
Приложение 1. Моделирование равномерно распределенного вектора в заданной области .....	39
Приложение 2. Метод Монте-Карло для оценки интегралов .....	39
Приложение 3. Моделирование нормально распределенного случайного вектора [11] .....	41
1). Моделирование двумерного нормального распределения .....	41
2). Моделирование многомерного нормального распределения .....	42
Приложение 4. Листинг программы. ....	44

## Введение

Наличие большого числа прикладных задач стимулировало развитие математической теории оптимального эксперимента.

В условиях дороговизны экспериментов или при проведении экспериментов, которые невозможно воспроизвести повторно, требуется предварительное планирование. Здесь важным для нас является понятие имитационного эксперимента, направленного на изучение реального процесса с последующим построением имитационной модели и получения реализаций некоторой случайной функции. Квалифицированное планирование приводит к сокращению времени, числа и стоимости измерений, в повышении точности восстановления неизвестной зависимости, в возрастании достоверности полученных выводов.

Во многих областях науки, например, таких как ядерная физика, спектроскопия, дистанционное зондирование атмосферы, регистрация и анализ изображений, анализ состава веществ, оптические методы индикации неоднородностей и загрязнений среды, цветовые измерения и т.д. возникают задачи восстановления распределений различной природы и оценивания их параметров [17,24,26,28,38].

Когда создана математическая модель и выбран метод планирования можно приступить к имитированию эксперимента с помощью компьютера, т. е. разработке и реализации алгоритма, возвращающего результаты, аналогичные измеренным данным. В том случае, когда мы имеем дело с задачей восстановления неизвестной функции по результатам измерения ее значений, требуется подобрать подходящую функцию, а затем вычислять ее значения и генерировать ошибку с помощью генератора случайных чисел.

Понятно, что применение методов планирования эксперимента может дать значительный экономический эффект, но отсутствие соответствующего оптимального плана может сделать экспериментальную программу полностью безрезультатной.

Основное внимание в этой работе уделено получению оптимального плана.

Планирование эксперимента предполагает наличие математической модели и критерия оптимальности. Сложный характер критериев требует специальных методов оптимизации. В диссертации рассматриваются два метода случайного поиска глобального экстремума: метод Метрополиса и ранее изученный метод случайного поиска глобального экстремума с “памятью”. Оба метода позволяют получать реализации случайного вектора с плотностью, пропорциональной целевой функции и далее находить моду рассматриваемой плотности и тем самым получать максимум исходной функции.

Магистерская диссертация является продолжением работы над бакалаврским дипломом, где находились оптимальные планы для линейной ре-

грессии с использованием случайного поиска.

Настоящая работа прошла апробацию на международных научных конференциях:

1. III международной конференции “Устойчивость и процессы управления”, 5-9 октября 2015 г.
2. III International Conference “Stability and Control Processes” in Memory of V.I.Zubov, 2015.
3. 48 международной научной конференции аспирантов и студентов “Процессы управления и устойчивость”, 3-6 апреля 2017 г.

## Постановка задачи

Требуется найти оптимальный план, используя подходящий критерий оптимальности.

Уравнение регрессии по отношению к физической задаче, связанной с выявлением отдельных радионуклидов в смеси двух радионуклидов, имеет вид

$$\eta(x, \theta) = \beta_1 e^{-\theta_1 x} + \beta_2 e^{-\theta_2 x}, \quad x \in [0, B] \subset R^1,$$

где  $x$  время. Для получения удовлетворительных результатов время меняется в промежутке  $x \in [0, B]$ , где константа  $B$  – должна быть не менее двух полупериодов распада наиболее долгоживущего компонента смеси  $T$ , то есть  $B \geq 2T$ . Линейные параметры  $\beta_1, \beta_2$  соответствуют количеству компонентов смеси и их вкладов в суммарную активность  $G_1 + G_2 = G_0$ . Нелинейные параметры  $\theta_1, \theta_2$  – определяют тип изотопа, так как они выражаются через постоянные величины полупериоды  $\theta_i = \frac{\ln 2}{T_i}$ ,  $i = 1, 2$ .

Задача очень сложная. Можно по-разному подходить к ней в зависимости от априорных данных относительно параметров.

Для нахождения параметров регрессии можно использовать нелинейный вариант метода наименьших квадратов [20,12]. Найти

$$\min_{\beta_1, \beta_2, \theta_1, \theta_2} \sum_{i=1}^N (G(x_i) - \eta(x_i, \beta_1, \beta_2, \theta_1, \theta_2))^2.$$

Как показывает практика, целевая функция является сложной и многоэкстремальной. Можно подобрать подходящие методы оптимизации, но наиболее достоверные результаты могут получиться при достаточно большом  $N$ , то есть большом количестве наблюдений. Остается очень важный вопрос: в каких точках разумнее делать эти наблюдения? Отсутствие оптимального плана может сделать экспериментальные наблюдения полностью безрезультатными. Кроме того, эти наблюдения и опасны, и дорогостоящие [1].

Если априори на основе опытных данных известны оценки для нелинейно входящих параметров  $\theta_1, \theta_2$ , то далее решать линейную задачу и находить параметры  $\beta_1, \beta_2$ , входящие линейно. Это решается в [24] и находятся оптимальные интервалы наблюдений каждого изотопа в смеси.

# Глава I. Обзор литературы

Методы планирования эксперимента, возникшие первоначально в связи со статистическими приложениями в медицине и сельском хозяйстве, получили в настоящее время широкое развитие и новые области применения.

Рональд Эйлмер Фишер, английский статистик, по праву считается основателем теории планирования эксперимента. В ряде работ, Фишер рассматривает задачи повышения урожайности сельскохозяйственной продукции и разрабатывает методы статистического анализа и планирования эксперимента для их решения [32].

Большой вклад в развитие теории планирования эксперимента внесли американские ученые Дж. Бокс и Дж. Кифер [29, 31]. Боксу принадлежат результаты по планированию регрессионного эксперимента для полиномиальных функций регрессии, конструированию критериев оптимальности. Кифер является создателем классической теории оптимального планирования регрессионного эксперимента.

Ряд направлений в теории планирования эксперимента возник в нашей стране благодаря А.Н. Колмогорову, Ю.В. Линнику и их ученикам [18,19]. Существенный вклад в развитие теории планирования эксперимента и ее приложений внесли ученые В.В. Федоров, С.М. Ермаков, Е.В. Седунов, В.М. Буре, В.Б. Мелас, А.А. Жиглявский [27, 10-12, 24, 2, 3, 21,22].

Будем рассматривать некоторую случайную функцию определенного вида с параметрами, которые следует оценить. Обозначим ее через  $\eta(\mathbf{x}, \boldsymbol{\theta})$ , где  $\mathbf{x}$  – заданный вектор из некоторого множества  $D \subset E^n$  и  $\boldsymbol{\theta} \in \Theta \subset E^m$  – вектор параметров.

В задачах слежения за движущимися объектами часто используется регрессионная функция вида

$$\eta(\mathbf{x}, \boldsymbol{\theta}) = \arctg \left[ \frac{x_1 - \theta_1}{x_2 - \theta_2} \right], \quad \mathbf{x} = (x_1, x_2), \quad \boldsymbol{\theta} = (\theta_1, \theta_2),$$

либо модель

$$\eta(x, \boldsymbol{\theta}) = \sum_{i=1}^l \frac{\theta_i}{(x - \theta_{i+1})^2 + \theta_{i+2}}, \quad \boldsymbol{\theta} = (\theta_1, \dots, \theta_{2l}).$$

В результате решения систем дифференциальных уравнений широко используется модель

$$\eta(x, \boldsymbol{\theta}) = \sum_{i=1}^l \theta_i \exp\{-\theta_{i+1}x\}, \quad \boldsymbol{\theta} = (\theta_1, \dots, \theta_{2l}).$$

Эта модель также используется в атомной физике. По найденным оценкам нелинейных параметров можно определять типы изотопов, а по оценкам линейных параметров – относительный вклад каждого изотопа в суммарную активность.

Остановимся подробнее на данном примере, так как предметом диссертации является оптимальное планирование эксперимента для нелинейной экспоненциальной регрессии.

### 1.1. Основной закон радиоактивного распада

Способность ядер самопроизвольно распадаться, испуская частицы, называется радиоактивностью. Радиоактивный распад - статистический процесс. Каждое радиоактивное ядро может распасться в любой момент, и закономерность наблюдается только в среднем, в случае распада достаточно большого количества ядер.

Каждое радиоактивное ядро распадается независимо от поведения всех других ядер, а потому общая скорость распада, т.е. число ядер, распадающихся в единицу времени (активность) пропорционально числу имеющихся радиоактивных ядер.

Самопроизвольные превращения радиоактивных ядер приводят к непрерывному уменьшению числа атомов (ядер) исходного радиоактивного изотопа и к образованию дочерних продуктов.

Радиоактивный распад относится к разряду вероятностных процессов, и к нему применимы методы статистического анализа.

Уравнение радиоактивного распада (основной закон радиоактивного распада в дифференциальной форме) имеет вид [1]:

$$\frac{dN}{dt} = -\lambda N, \quad (1.1)$$

где  $N$  – число атомов, не претерпевших распад к моменту времени  $t$ ,  $\lambda$  – константа. Величину  $G = \lambda N$  называют *радиоактивностью радионуклида*.

Коэффициент пропорциональности  $\lambda$  называется константой (постоянной) радиоактивного распада (радиоактивной постоянной) и он равен *вероятности распада каждого отдельного ядра за единицу времени*.

Константа  $\lambda$  характеризует неустойчивость ядер радиоактивного изотопа.

Из равенства

$$\lambda = -\frac{1}{N} \frac{dN}{dt}, \quad (1.2)$$

которое получается из (1.1), очевидно, что постоянная распада  $\lambda$  численно равна доле атомов  $\frac{dN}{N}$  распадающихся в единицу времени, при условии, что единица времени достаточно мала по сравнению с периодом полураспада.

Одно из свойств радиоактивного распада заключается в том, что за равные промежутки времени подвергается распаду постоянная часть от общего количества имеющихся в данный момент атомов радиоактивного изотопа. На рис. 1.1 величина  $T$  равна времени, за которое распалось половина имеющихся ядер.

Интегрирование уравнения (1.1) при условии, что в начальный момент времени  $t=0$  количество радиоактивных ядер составляет  $N_0$ , дает следующее соотношение

$$N = N_0 e^{-\lambda t} . \quad (1.3)$$

По формуле (1.3) получаем число распавшихся ядер ко времени  $t$ . В равные промежутки времени всегда распадаются равные части имеющихся радиоактивных атомов. В качестве меры устойчивости радиоактивного нуклида используют период полураспада  $T$ .

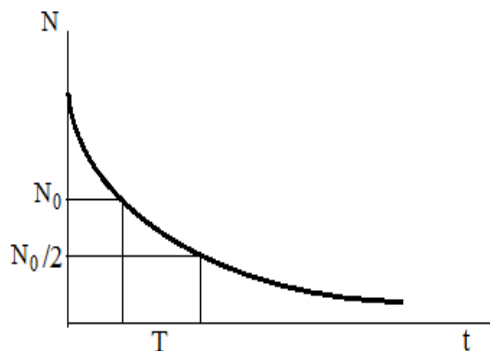


Рис. 1.1. Кривая радиоактивного распада

Нетрудно получить, что

$$T = \frac{\ln 2}{\lambda} = \frac{0.69315}{\lambda} . \quad (1.4)$$

## 1.2. Распад смеси радионуклидов

Если несколько радиоактивных веществ смешать вместе, то наблюдаемая полная активность представляет собой сумму активностей всех веществ

$$G(t) = \sum_{i=1}^n G_i e^{-\lambda_i t} , \quad (1.5)$$

где  $\lambda_i$  – постоянная распада  $i$ -го радионуклида,  $n$  – число радионуклидов в смеси.



Обычно представляют (1.5) в следующем виде

$$G(t) = G_0 \sum_{i=1}^n p_i e^{-\lambda_i t}, \quad (1.6)$$

где  $G_0 = \sum_{i=1}^n G_i$ ,  $p_i = \frac{G_i}{G_0}$ ,  $\sum_{i=1}^n p_i = 1$ ,  $G_i = G_0 p_i$ .

Исходя из законов распада, можно интерпретировать задачу нахождения оценки параметров нелинейной регрессии, которая рассматривается в настоящей работе, следующим образом. Оценки нелинейных параметров, связанные с полупериодами распада изотопов, дают возможность определять типы изотопов, а оценки линейных параметров представляют относительный вклад каждого изотопа в суммарную активность.

Найденный оптимальный план дает возможность разумно спланировать эксперимент. В данном случае указывает определенные моменты времени, при которых имеет смысл проводить эксперименты.

### 1.3. Математическая постановка задачи

Запишем функцию регрессии (1.5) при  $n = 2$  в следующем виде

$$\eta(x, \beta_1, \beta_2, \theta_1, \theta_2) = \beta_1 e^{-\theta_1 x} + \beta_2 e^{-\theta_2 x}, \quad 0 \leq x \leq B, \quad (1.7)$$

где  $B$  – некоторая заданная константа. В (1.7) выделены линейные параметры через  $\beta_1, \beta_2$ , а нелинейные параметры обозначены как  $\theta_1, \theta_2$ . Требуется найти оптимальный план, используя подходящий критерий оптимальности.

В диссертации предлагается находить оптимальный план при наличии априорной информации для нелинейных параметров. Минимальное количество точек плана для рассматриваемой задачи равно 4 (насыщенный оптимальный план). В качестве критерия оптимальности выбрана функция  $A(x_1, x_2, x_3, x_4)$  от точек плана, максимум которой соответствует минимуму объема эллипсоида рассеяния ошибок при оценке всех параметров функции регрессии.

## Глава II. Оптимальный план для регрессионного эксперимента

Термин “регрессия” возник в исследованиях английского естествоиспытателя Френсиса Гальтона (1822-1911) по биологической наследственности [2].

### 2.1. Классическая схема регрессии

Рассмотрим нелинейную регрессионную модель, которая представима некоторой вещественной функцией  $\eta(x, \theta)$ ,  $x \in X \subset R^n$ ,  $\theta \in \Xi \subset R^m$ . Эта функция известна с точностью до вектора параметров  $\theta = (\theta_1, \dots, \theta_m)$ , причем хотя бы один из этих параметров входит нелинейно в эту модель. Множество  $X$  является компактом и называется множеством планирования.

Опишем классическую схему регрессионного эксперимента. Пусть в точках  $x_j, j = \overline{1, N}$  из множества планирования  $X$  наблюдаются значения случайных величин  $y_j$ , представимых в виде

$$y_j = \eta(x_j, \theta) + \varepsilon_j, \quad j = \overline{1, N}, \quad (2.1)$$

где  $\varepsilon_1, \dots, \varepsilon_N$  – случайные ошибки измерений. Предполагаем,  $\varepsilon_j$  центрированными и некоррелированными:  $M\varepsilon_j = 0$ ,  $M\varepsilon_i\varepsilon_j = 0$ ,  $j \neq i$ ,  $D\varepsilon_i = M\varepsilon_i^2 = \sigma^2$ .

Для оценивания неизвестных параметров  $\theta = (\theta_1, \dots, \theta_m)$  используется один из наиболее распространенных методов оценивания параметров нелинейных регрессий – метод наименьших квадратов (МНК):

$$\hat{\theta}_N = \arg \min_{\tilde{\theta} \in \Xi} \sum_{i=1}^N (y_i - \eta(x_i, \tilde{\theta}))^2. \quad (2.2)$$

Для введенной схемы регрессионного эксперимента (2.1) при достаточно слабых предположениях с вероятностью 1 при  $N \rightarrow \infty$  верны следующие утверждения [12,15,22,35]:

$$\hat{\theta}_N \rightarrow \theta, \quad (2.3)$$

$$D(\hat{\theta}_N - \theta) / \sqrt{N} \rightarrow \sigma^2 M^{-1}(x, \theta), \quad (2.4)$$

где  $M(x, \theta)$  – информационная матрица размерности  $m \times m$ . Она находится через матрицу  $\varphi(x, \theta) = \left( \frac{\partial \eta(x_j, \theta)}{\partial \theta_1}, \dots, \frac{\partial \eta(x_j, \theta)}{\partial \theta_m} \right)_{j=1}^N$  следующим образом:

$$M(x, \theta) = \varphi^T(x, \theta)\varphi(x, \theta). \quad (2.5)$$

Таким образом, на основании (2.3), (2.4) можно сказать, что последовательность оценок (МНК)  $\hat{\theta}_N$  является состоятельной, а последовательность распределений случайных векторов  $(\hat{\theta}_N - \theta)/\sqrt{N}$  сходится к нормальному распределению с нулевым вектором средних и дисперсионной матрицей  $\sigma^2 M^{-1}(x, \theta)$ .

В связи с этим в качестве критерия оптимальности для нахождения оптимального плана могут быть выбраны те же критерии, что и для линейной регрессии, например, критерий  $D$ -оптимальности: максимизация определителя информационной матрицы

$$\det M(x, \theta) \rightarrow \sup_{x \in X}. \quad (2.6)$$

В нелинейном случае информационная матрица зависит от вектора истинных значений параметров  $\theta$ .

В случае нормально распределенных ошибок, как в классической схеме регрессии, этот критерий соответствует требованию минимизации объема доверительного эллипсоида с любым фиксированным уровнем доверия.

Решение задачи (1.6) при некотором фиксированном значении  $\theta$  называется задачей поиска *локально оптимального плана*. Такой подход к решению этой задачи впервые рассматривался в работе [30], а также в работе [32].

## 2.2. Нелинейная экспоненциальная регрессия

Рассмотрим нелинейную экспоненциальную регрессию

$$\eta(x, \theta) = \beta_1 e^{-\theta_1 x} + \beta_2 e^{-\theta_2 x}, \quad x \in [0, B] \subset R^1, \quad \beta_1, \beta_2 \neq 0, \quad \theta_2 > \theta_1 > 0, \quad (2.7)$$

с вектором параметров  $\theta = (\beta_1, \theta_1, \beta_2, \theta_2)$ . В (2.7)  $B$  – некоторая заданная величина,  $n = 1$ . Как мы видим, только параметры  $\theta_1, \theta_2$  входят в (2.7) нелинейно.

Будем рассматривать минимальное число узлов  $N = 4$  (насыщенный план). Определитель информационной матрицы для функции регрессии (2.7) равен [20],

$$\det(M(x, \theta)) = \det(\varphi^T(x, \theta))(\varphi(x, \theta)) = (\det(\varphi^T(x, \theta)))^2,$$

где

$$\det(\varphi^T(x, \theta)) = \beta_1 \beta_2 \begin{vmatrix} e^{-\theta_1 x_1} & e^{-\theta_1 x_2} & e^{-\theta_1 x_3} & e^{-\theta_1 x_4} \\ -x_1 e^{-\theta_1 x_1} & -x_2 e^{-\theta_1 x_2} & -x_3 e^{-\theta_1 x_3} & -x_4 e^{-\theta_1 x_4} \\ e^{-\theta_2 x_1} & e^{-\theta_2 x_2} & e^{-\theta_2 x_3} & e^{-\theta_2 x_4} \\ -x_1 e^{-\theta_2 x_1} & -x_2 e^{-\theta_2 x_2} & -x_3 e^{-\theta_2 x_3} & -x_4 e^{-\theta_2 x_4} \end{vmatrix}.$$

Следовательно, определитель информационной матрицы равен

$$\det(M(x, \theta)) = \beta_1^2 \beta_2^2 A(x, \theta_1, \theta_2), \quad (2.8)$$

где

$$A(x_1, x_2, x_3, x_4, \theta_1, \theta_2) = \begin{vmatrix} e^{-\theta_1 x_1} & e^{-\theta_1 x_2} & e^{-\theta_1 x_3} & e^{-\theta_1 x_4} \\ -x_1 e^{-\theta_1 x_1} & -x_2 e^{-\theta_1 x_2} & -x_3 e^{-\theta_1 x_3} & -x_4 e^{-\theta_1 x_4} \\ e^{-\theta_2 x_1} & e^{-\theta_2 x_2} & e^{-\theta_2 x_3} & e^{-\theta_2 x_4} \\ -x_1 e^{-\theta_2 x_1} & -x_2 e^{-\theta_2 x_2} & -x_3 e^{-\theta_2 x_3} & -x_4 e^{-\theta_2 x_4} \end{vmatrix}^2 \quad (2.9)$$

функция, зависящая от переменных  $x_1, x_2, x_3, x_4$  (точки множества планирования) с входящими в нее нелинейными параметрами  $\theta_1, \theta_2$ . Линейные параметры  $\beta_1, \beta_2$  являются множителями. Для определенности будем полагать  $\beta_1 = \beta_2 = 1$ .

### 2.3. Локально D-оптимальные планы для экспоненциальной регрессии

Задача нахождения локально оптимального плана для экспоненциальной регрессии сводится к нахождению максимума (2.9) по точкам из множества планирования с фиксированными нелинейными параметрами  $\theta_{1,ист}, \theta_{2,ист}$ , которые будем называть истинными,

$$\max_{x_1, x_2, x_3, x_4 \in X} A(x_1, x_2, x_3, x_4, \theta_{1,ист}, \theta_{2,ист}). \quad (2.10)$$

Будем предполагать, что  $\theta_{1,ист} > \theta_{2,ист}$ , так как если  $\theta_{1,ист} = \theta_{2,ист}$ , то определитель в (2.9) будет тождественно равен нулю из-за совпадения двух строк. Можно рассмотреть следующие варианты [22]:

$$\theta_{1,ист} = 1 + \Delta_i, \quad \theta_{2,ист} = 1 - \Delta_i, \quad \Delta_i = 0.1 * i, \quad i = \overline{1,9}. \quad (2.11)$$

Далее в главе II и III для решения задачи (2.10) предлагается два метода поиска глобального экстремума. Оба метода основаны на случайном поиске глобального экстремума и используют реализации случайных векторов, рас-

пределенных с плотностью, связанной с целевой функцией. В результате поиск максимума функции  $A(x_1, x_2, x_3, x_4, \theta_{1,ист}, \theta_{2,ист})$  можно трактовать как нахождение моды плотности распределения вероятностей [36].

Первый из методов является итерационным и опирается на вычисление элементов матрицы ковариаций. Эта матрица от итерации к итерации передает свойства целевой функции и позволяет сосредоточить область поиска в окрестности найденной приближенной экстремальной точки. Этот метод требует знания моделирования нормально распределенных векторов.

Второй метод – это метод Метрополиса. Он используется для моделирования случайных векторов, распределенных с плотностью, связанной с целевой функцией, и основан на моделировании цепей Маркова.

## Глава III. Метод случайного поиска глобального экстремума на основе оценивания элементов матрицы ковариаций

При решении достаточно сложных многоэкстремальных задач часто используются методы случайного поиска глобального экстремума [4,5,11,13,14,16,23,36]. Для методов случайного поиска вид множества оптимизации  $D$  имеет не слишком большое значение. Важно лишь, чтобы оно было не слишком экзотическим. Ограничения, возникающие при моделировании случайных векторов в заданной области  $D$ , могут учитываться простейшим образом: если для полученной пробной точки не выполняются ограничения, она отбрасывается и моделируется новая точка.

Предполагается, что требуется найти вектор  $\hat{\mathbf{x}} \in D$ , удовлетворяющий условию

$$f(\hat{\mathbf{x}}) = \max_{\mathbf{x} \in D} f(\mathbf{x}). \quad (3.1)$$

Заметим, что задачи поиска минимума функции сводятся к задаче поиска максимума путем замены знака целевой функции  $f$  или можно рассматривать обратную функцию  $\frac{1}{f}$ .

В простейших алгоритмах случайного поиска организовывается “просмотр” всей области  $D$ . В этом случае используются случайные равномерно распределенные вектора. Далее можно использовать информацию о целевой функции  $f$  и организовать неравномерное моделирование. Тогда будем иметь возможность получать пробные точки в подмножестве  $D$ , где находится наиболее перспективная точка, то есть точка, в которой значение  $f$  наибольшее. “Обучение” алгоритма вводится тем, что с помощью матрицы ковариаций, которая от итерации к итерации несет в себе более точные свойства функции  $f$ , корректируется область поиска экстремума.

### 3.1. Простейший случайный поиск (слепой поиск)

Простейший случайный поиск глобального экстремума является вспомогательным для случайного поиска с “памятью”.

Предположим, что функция  $f(x)$  кусочно-непрерывна в некоторой области  $D \subset E^n$   $n$ -мерного евклидова пространства и требуется найти приближённо точку  $\hat{x}$ , удовлетворяющую (3.1).

Приближённых методов отыскания максимума функции очень много. Однако большая часть из них является локальными методами, сходимость которых гарантируется в достаточно малой окрестности экстремальной точки.

Простейший случайный поиск представляет собой метод приближённого поиска глобального максимума и основан на просмотре многомерной области  $D$  с помощью равномерно распределённых в этой области случайных точек.

### Алгоритм простейшего случайного поиска

1). Моделируем  $N$  точек

$$\xi^{(i)} = (\xi_1^{(i)}, \dots, \xi_n^{(i)}), \quad i = \overline{1, N}, \quad (3.2)$$

равномерно распределённых в области  $D$ .

2). Получаем приближенное максимального значения функции.

Положим

$$s_1 = f(\xi^{(1)}),$$

$$s_i = \begin{cases} f(\xi^{(i)}), & f(\xi^{(i)}) > s_{i-1}, \\ s_{i-1}, & f(\xi^{(i)}) \leq s_{i-1} \end{cases}, \quad i = \overline{2, N}. \quad (3.3)$$

Тогда

$$f(\hat{x}) \approx s_N.$$

### 3.2. Доказательство сходимости простейшего случайного поиска

Введём  $\varepsilon$ -окрестность точки  $\hat{x}$  —  $S_\varepsilon(\hat{x})$  и событие  $A$ , состоящее в том, что  $\xi^{(k)} \in S_\varepsilon(\hat{x})$ . Вероятность этого события равна  $P(A) = \frac{|S_\varepsilon(\hat{x})|}{|D|} = p$ , где  $|D|$  означает объём области  $D$ . Пусть  $q = 1 - p$  ( $q < 1$ ). Таким образом, каждая точка (3.2) с вероятностью  $p$  попадает в  $S_\varepsilon(\hat{x})$  и с вероятностью  $q$  не попадает в  $S_\varepsilon(\hat{x})$  (распределение Бернулли). Найдём вероятность  $P_N$  того, что хотя бы одна из  $N$  пробных точек попадет в  $S_\varepsilon(\hat{x})$

$$P_N = P \left\{ \begin{array}{l} \text{хотя бы одна из } N \text{ точек} \\ \text{принадлежит } S_\varepsilon(\hat{x}) \end{array} \quad \xi^{(k)} \right\} = 1 - q^N.$$

Откуда имеем: если  $N \rightarrow \infty$ , то  $P_N \rightarrow 1$ .

Таким образом доказали следующую теорему.

**Теорема 3.1.** Вероятность того, что хотя бы одна пробная точка (3.2) из  $N$  испытаний попадёт в  $\varepsilon$ -окрестность точки  $\hat{x}$ , доставляющей максимум функции  $f(x)$ , стремится к 1 при  $N \rightarrow \infty$ .

Эта теорема может служить доказательством сходимости простейшего случайного поиска.

### 3.3. Оценивание числа испытаний

Пусть  $\frac{|S_\varepsilon(\hat{x})|}{|D|} = p = \varepsilon, q = 1 - \varepsilon$ .

Запишем условие того, что вероятность попадания хотя бы одной точки в окрестность  $S_\varepsilon(\hat{x})$  не меньше  $1 - \varepsilon$ :  $P_N \geq 1 - \varepsilon$  или  $1 - (1 - \varepsilon)^N \geq 1 - \varepsilon$ . Тогда  $(1 - \varepsilon)^N \leq \varepsilon$ . Откуда  $N \geq \frac{\ln \varepsilon}{\ln(1 - \varepsilon)}$ . Следовательно,

$$N \approx \left[ \frac{\ln \varepsilon}{\ln(1 - \varepsilon)} \right], \quad (3.4)$$

где  $[x]$  – целая часть  $x$ .

Зависимость числа испытаний от заданной точности  $\varepsilon$  представлена в следующей таблице

Таблица 3.1

$\varepsilon$	0.5	0.2	0.1	0.01	0.001
$N$	1	8	22	458	6903

### 3.4. Случайный поиск с “памятью”

**Идея метода.** Пусть решается задача (3.1). Очевидно, что если организовать поиск максима функции, используя неравномерное распределение, например, в качестве плотности распределения взять плотность распределения вероятностей

$$p(x) = \frac{f(x)}{\int_D f(x) dx}, \quad (3.5)$$

то пробные точки будут чаще попадать в окрестность  $\hat{x}$ . Более того, так как

$$\lim_{m \rightarrow \infty} f^m(x) = \delta(x - \hat{x}),$$

где  $\delta(x - \hat{x})$  – функция Дирака, сосредоточенная в  $\hat{x}$ , то математическое ожидание случайной величины  $\xi$ , распределённой с плотностью



$$p_m(x) = \frac{f^m(x)}{\int_D f^m(x) dx}, \quad (3.6)$$

будет стремиться  $\hat{x}$  при  $m \rightarrow \infty$ . Действительно,

$$M\xi = \int_D x p_m(x) dx \xrightarrow{m \rightarrow \infty} \int_D x \delta(x - \hat{x}) dx = \hat{x}.$$

Мы учли ещё и следующее свойство функции Дирака

$$\int_D f^m(x) dx \xrightarrow{m \rightarrow \infty} \int_D \delta(x - \hat{x}) dx = 1.$$

### 3.5. Введение неравномерного распределения

Однако моделирование (3.6) может быть затруднительно, поэтому вместо  $p_m(x)$  предлагается рассматривать, например, плотность [5,13,14]

$$\bar{p}_m(x) = G_m f(x) N(\bar{x}^{(m-1)}, \bar{C}^{(m-1)}), \quad m = \overline{1, M}, \quad (3.7)$$

где

$$G_m = \frac{1}{\int_D f(x) N(\bar{x}^{(m-1)}, \bar{C}^{(m-1)}) dx} - \text{константа нормировки, а функция}$$

$$\begin{aligned} N(\bar{x}^{(m-1)}, \bar{C}^{(m-1)}) &= \\ &= \frac{\det \bar{C}^{(m-1)}{}^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2}(x - \bar{x}^{(m-1)})^T (\bar{C}^{(m-1)})^{-1} (x - \bar{x}^{(m-1)})\right\} \end{aligned} \quad (3.8)$$

является нормальной плотностью с параметрами:  $\bar{x}^{(m-1)}$  – вектор, являющийся приближением точного значения  $\hat{x}$ ,  $\bar{C}^{(m-1)}$  – матрицей ковариаций. Матрица  $\bar{C}^{(m-1)}$  является приближением матрицы  $C^{(m-1)}$ , элементы которой вычисляются по формулам

$$\begin{aligned} C_{i,j}^{(m-1)} &= \iint_D x_i x_j p_m(x) dx - \iint_D x_i p_m(x) dx \iint_D x_j p_m(x) dx = \\ &= G_m \iint_D x_i x_j f(x) N(\bar{x}^{(m-1)}, C^{(m-1)}) dx - \\ &- G_m \iint_D x_i f(x) N(\bar{x}^{(m-1)}, C^{(m-1)}) dx \times G_m \iint_D x_j f(x) N(\bar{x}^{(m-1)}, C^{(m-1)}) dx, \quad m = \overline{1, M}, \end{aligned} \quad (3.9)$$

то есть матрицей ковариаций для случайной величины, распределенной с плотностью вероятностей (3.8).

Нормальная плотность, входящая в  $\bar{p}_m(x)$ , а именно, матрица ковариаций, несёт в себе информацию о функции  $f(x)$ . В связи с этим метод имеет название с “памятью”.

Нормальное распределение позволяет получать на  $j$ -ой итерации  $j = \overline{1, M}$ , больше реализаций в эллипсоиде рассеяния

$$(x - \bar{x}^{(j-1)}) (\bar{C}^{(j-1)})^{-1} (x - \bar{x}^{(j-1)})^T = 1 \quad (3.10)$$

с центром в точке  $\bar{x}^{(j-1)}$  и матрицей  $\bar{C}^{(j-1)}$ . Можно сказать, что происходит “обучение” поиска. В каждой итерации  $j$  область поиска уточняется. Эллипсоид рассеяния (3.10) имеет центр  $\bar{x}^{(j-1)}$  – точку, которая считается наиболее перспективной на предыдущей итерации. В следующей итерации эта точка может быть только улучшена в силу алгоритма метода (3.3).

Интегралы, входящие в (3.9), можно оценить по методу Монте-Карло [6,11]. Поэтому оценки элементов матрицы  $C^{(j-1)}$  будем получать по следующим формулам

$$\begin{aligned} c_{i,j}^{(m-1)} &\approx \bar{c}_{i,j}^{(m-1)} = \\ &= \frac{\sum_{k=1}^{N_m} x_i^{(k)} x_j^{(k)} f(x^{(k)})}{\sum_{k=1}^{N_m} f(x^{(k)})} - \frac{\sum_{k=1}^{N_m} x_i^{(k)} f(x^{(k)}) \sum_{k=1}^{N_m} x_j^{(k)} f(x^{(k)})}{\left( \sum_{k=1}^{N_m} f(x^{(k)}) \right)^2}, \end{aligned} \quad (3.11)$$

где  $x^{(k)}$ ,  $k = \overline{1, N_m}$ , – реализации случайного вектора, имеющего плотность (3.8), т.е. нормально распределенного.

### 3.6. Алгоритм метода случайного поиска с “памятью”

### **Предварительная итерация $m = 1$ .**

На этой итерации используется простейший случайный поиск экстремума функции.

1). Получаем последовательность векторов

$$\xi^{(k)} = (\xi_1^{(k)}, \dots, \xi_n^{(k)}), \quad k = \overline{1, N_0}, \quad (3.12)$$

являющихся реализациями случайного вектора  $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ , равномерно распределённого в области  $D$  изменения функции  $f(x)$  (алгоритм описан в приложении 1). Далее по (3.3) находим точку  $\bar{x}^{(0)}$ , являющуюся приближением для экстремальной точки  $\hat{x}$ .

2). Параллельно с поиском  $\bar{x}^{(0)}$  оценим матрицу ковариаций  $C^{(0)}$  для случайной величины  $\xi$ , имеющую плотность распределения

$$p_0(x) = G_0 f(x), \quad G_0 = \frac{1}{\iint_D f(x) dx}. \quad (3.13)$$

Элементы матрицы ковариаций  $C^{(0)}$  находятся по формулам

$$c_{i,j}^{(0)} = \iint_D x_i x_j p_0(x) dx - \iint_D x_i p_0(x) dx \iint_D x_j p_0(x) dx, \quad i, j = \overline{1, n}, \quad (3.14)$$

а оценку этих элементов проведём по методу Монте-Карло по формулам П.2.3(приложение 2) ( $m = 1$  и  $x^{(k)}$ ,  $k = \overline{1, N_0}$ , есть реализации (3.12)). Вид оценок элементов матрицы ковариаций (3.14) совпадает с (3.11). Но в этом случае  $x^{(k)}$ ,  $k = \overline{1, N_0}$ , равномерно распределены в области интегрирования.

### **В). $m$ -я итерация ( $m = \overline{2, M}$ )**

На этих итерациях моделируем точки по формулам П.3.9 (Приложение 3)

$$\xi^{(i)} = (\xi_1^{(i)}, \dots, \xi_n^{(i)}), \quad i = \overline{1, N_m}, \quad (3.15)$$

которые являются реализациями случайной нормально распределённой величины, т.е. с плотностью (3.8). Это можно выполнить, так как параметры нормального распределения  $\bar{x}^{(m-1)}$  и  $\bar{C}^{(m-1)}$  уже известны из предыдущих итераций, начиная с  $m = 1$ .

Далее аналогично проделываем следующую  $(m+1)$ -ю итерацию, пока число итераций будет меньше заданного  $M$ .

**Замечание.** С увеличением числа итераций полуоси эллипсоида рассеяния уменьшаются, что может сильно снизить вероятность попадания в  $S_\varepsilon(\hat{x})$  – окрестность экстремальной точки.

Обозначим через  $P_M$  вероятность того, что хотя бы одна точка при  $M$  итерациях попадёт в  $S_\varepsilon(\hat{x})$ . Пусть  $\lambda_m$  – наименьшее собственное число (через которое определяется минимальная полуось эллипсоида рассеяния) матрицы ковариаций  $\bar{C}^{(m)}$ .

**Теорема 3.2 [36].** Если  $\lambda_m = O(\ln^{-u} m)$  при  $0 < u < 1$ , то  $P_M \xrightarrow{M \rightarrow \infty} 1$ .

Доказательство.

$$P_M = 1 - q_1 q_2 \cdots q_M = 1 - \prod_{m=1}^M (1 - p_m), \quad (3.16)$$

где

$$p_m = \frac{1}{|D|(2\pi)^{\frac{n}{2}} \sqrt{\det \bar{C}_m} S_\varepsilon(\hat{x})} \int e^{-\frac{1}{2}(x-\bar{x}^{(m)})^T \bar{C}^{(m)}(x-\bar{x}^{(m)})} dx \quad (3.17)$$

– вероятность попадания в  $S_\varepsilon(\hat{x})$  на  $m$ -ой итерации.

Используя некоторые преобразования, получим оценку

Тогда

$$p_m > \frac{1}{|D|(2\pi)^{\frac{n}{2}} \sqrt{d_{1,m} \cdots d_{n,m}}} \times \int_{S'_\varepsilon} e^{-\frac{1}{2}\left(\frac{y_1^2}{d_{1,m}} + \cdots + \frac{y_n^2}{d_{n,m}}\right)} dy > B \frac{1}{(\mu_m)^n} e^{-\frac{n\gamma}{\lambda_m}}. \quad (3.18)$$

Здесь величины  $B$  и  $\gamma$  не зависят от  $m$ .

Пусть минимальное и максимальное собственные числа матриц  $\bar{C}^{(m)}$ ,  $m = \overline{1, M}$ , подчиняются следующим закономерностям:

$$\lambda_m = \min_{i=1, n} d_{i,m} = \ln^{-u} m, \quad (3.19)$$

$$\mu_m = \max_{i=1, n} d_{i,m} = \ln^{-v} m, \quad (3.20)$$

где  $u, v = const$  и  $0 < v < u$ .

Очевидно, что тогда

$$p_m > B \frac{\ln^{nv} m}{e^{m \ln^u m}} e^{-\frac{n\gamma}{\lambda_m}} = B \frac{\ln^{nv} m}{e^{\ln m \gamma \ln^{u-1} m}} = B \frac{\ln^{nv} m}{m^{\gamma \ln^{u-1} m}}. \quad (3.21)$$

Рассмотрим

$$\ln \prod_{m=1}^M (1 - p_m) = \sum_{m=1}^M \ln(1 - p_m) \leq - \sum_{m=1}^M p_m \stackrel{(3.21)}{<} -B \sum_{m=1}^M \frac{\ln^{\eta} m}{m^{\eta \ln^{u-1} m}} \xrightarrow{u < 1, M \rightarrow \infty} -\infty.$$

Из этого следует, что

$$\prod_{m=1}^M (1 - p_m) \xrightarrow{u < 1, M \rightarrow \infty} 0.$$

Это означает, что

$$P_M \xrightarrow{u < 1, M \rightarrow \infty} 1.$$

Таким образом, при определённой скорости уменьшения минимальной полуоси эллипсоида рассеяния, случайный метод с памятью сходится по вероятности.

Величина площади эллипсоида рассеяния определяется собственными числами матрицы  $\bar{C}^{(m-1)}$ .

Итерация повторяется многократно ( $m = \overline{2, M}$ ). На каждой итерации заново оцениваются приближенный максимум  $\bar{x}^{(m-1)}$  и находится  $\bar{C}^{(m-1)}$  – оценка матрицы ковариаций.

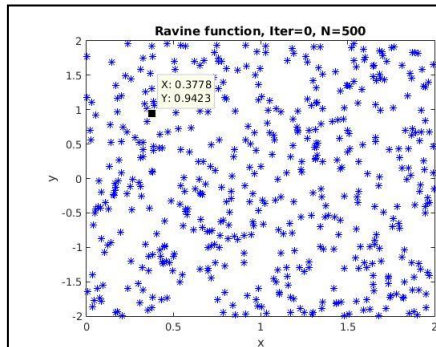


Рис. 3.2. Равномерное распределение

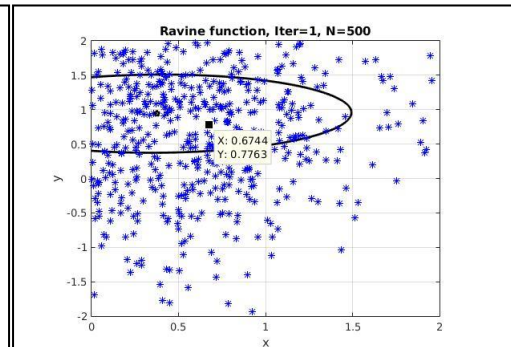


Рис. 3.3. Итерация 1

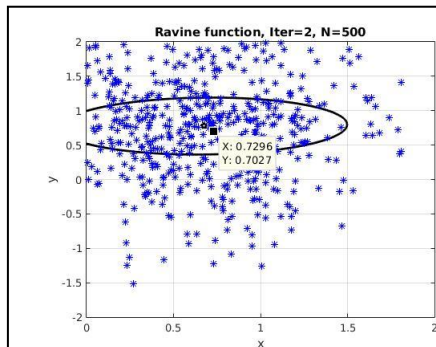


Рис. 3.4. Итерация 2

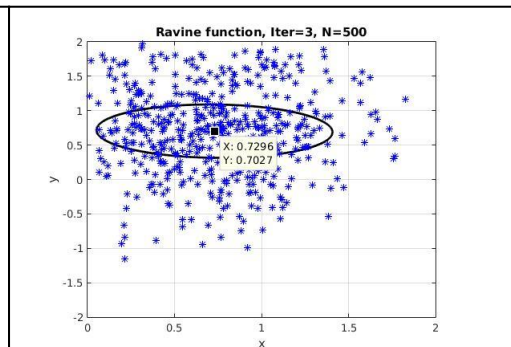


Рис. 3.5. Итерация 3

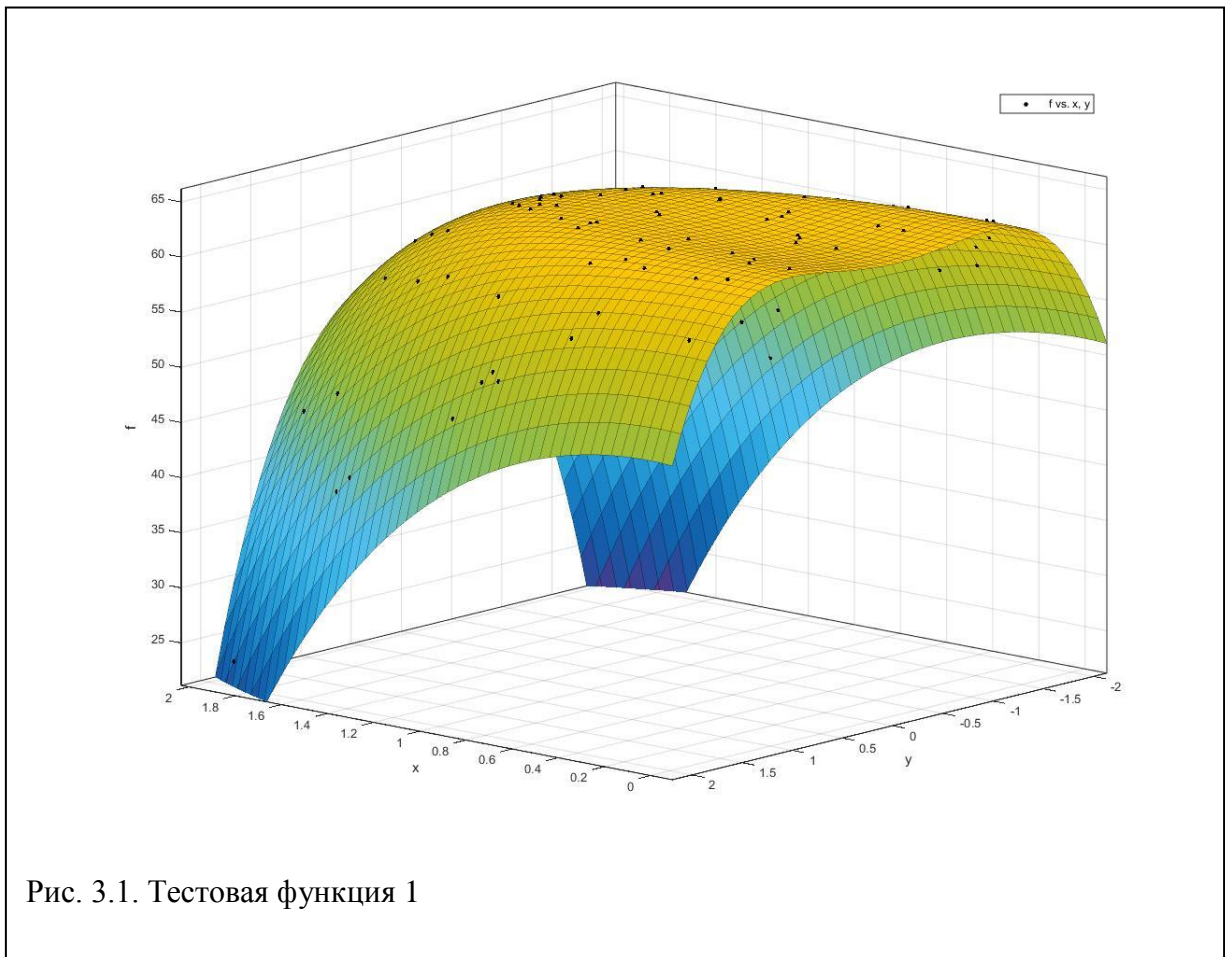
*Замечание 1.* При использовании нормального распределения вектора последовательности реализаций (3.15) чаще попадают в эллипсоид рассеяния (3.10), однако “хвосты” нормального распределения позволяют просматри-

вать всю область  $D$ . Таким образом, точка глобального экстремума не должна быть потеряна.

*Замечание 2.* Следует обратить внимание на то, что плотность (3.7) не моделируется, а находится матрица ковариаций (3.9) для случайного вектора, распределенного с этой плотностью. При оценивании элементов этой матрицы ковариаций по методу Монте-Карло требуются реализации случайного нормально распределенного вектора.

На вспомогательной итерации также плотность (3.5) не моделируется, а вычисляется матрица ковариаций для случайной величины, распределенной с плотностью (3.5). При оценке элементов этой матрицы по методу Монте-Карло моделируется равномерно распределенный вектор в области  $D$ .

Таким образом, удалось избежать моделирования сложных плотностей (3.5) и (3.7), хотя эти плотности участвуют в формировании неравномерного распределения.



В режиме тестирования программа ищет экстремум для т.н. овражной функции (тестовая функция 1, рис. 3.1)

$$f(x) = A - \left( \sum_{i=1}^n x_i^2 - a^2 \right)^2 + \alpha \sum_{i=1}^n x_i,$$

с параметрами  $A = 16n^2$ ,  $\alpha = 0.1$ ,  $a = 1$ . Область определения есть параллелепипед

$$D = \{0 \leq x_1 \leq 2, -2 \leq x_i \leq 2, i = \overline{1, n}\}.$$

Известно, что эта функция имеет наибольшее значение  $f \approx A$  при  $x_i \approx \frac{1}{\sqrt{n}}$ .

Результат работы программы тем точнее, чем больше выбирается число испытаний  $N$ . Для рассматриваемой функции Тест1 при  $n=2$  получено практически точное совпадение с ожидаемым экстремальным значением при  $x_i \approx 0.7$  на 3-ей итерации рис. 3.4.

Вторая тестовая функция (рис.3.6)

$$f(x) = \left| \sum_{i=1}^2 x_i^2 - \sum_{i=1}^2 \cos(18x_i) \right|,$$

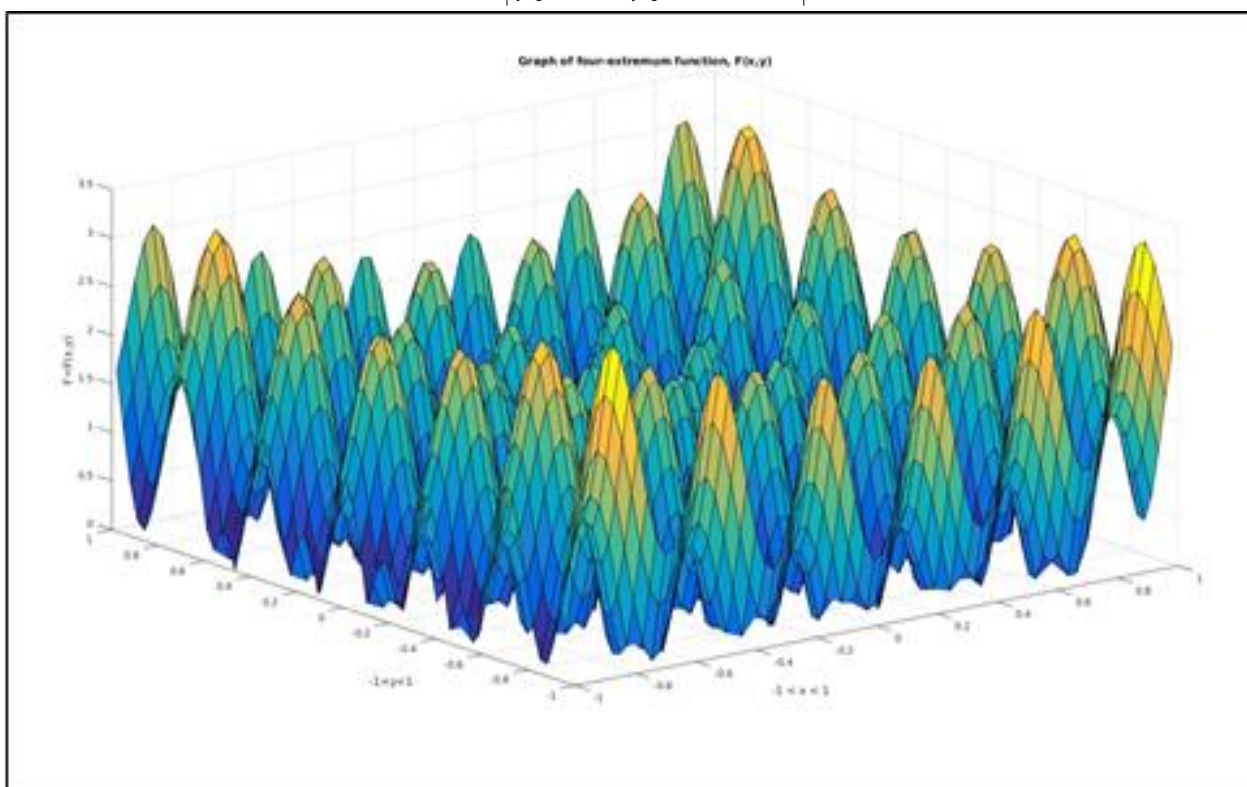


Рис. 3.6 Вторая тестовая функция имеет 4 глобальных экстремума и множество локальных в области  $D = \{-1 \leq x_i \leq 1, i = \overline{1, 2}\}$ .

Известно, что в точках

$$x^{(1)} = (-1, -1), \quad x^{(2)} = (-1, 1), \quad x^{(3)} = (1, -1), \quad x^{(4)} = (1, 1),$$

функция имеет наибольшее значение  $f_{\max}(x^{(i)}) \approx 3.5, i = \overline{1, 4}$ .

Результаты расчетов с использованием программы приведены на рис. 3.8-3.10.

При первом запуске программа быстро находит один из экстремумов, а при последующих запусках находит и другие максимумы.

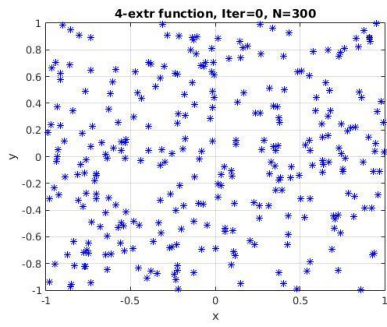


Рис 3.7 Равномерное распределение

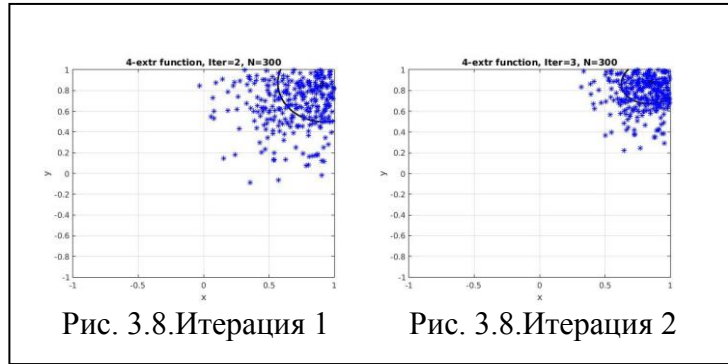


Рис. 3.8.Итерация 1

Рис. 3.8.Итерация 2

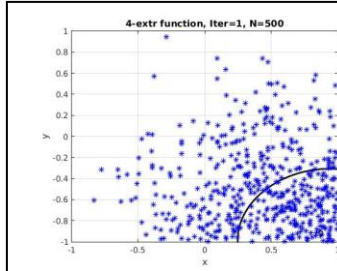


Рис. 3.9. Итерация 1

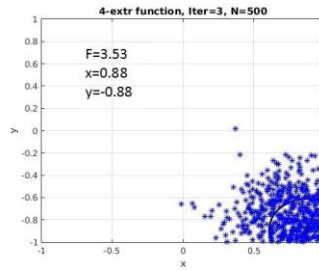


Рис. 3.10. Итерация 2



## Глава IV. Метод Метрополиса для нахождения моды плотности распределения случайных величин

Пусть имеется функция  $f(x)$ , заданная в области  $D \subset E^n$   $n$ -мерного евклидова пространства. Объем  $D$  предполагается ограниченным, а  $f(x) \geq 0, x \in D$ . Требуется найти

$$f^* = f(x^*) = \sup_D f(x). \quad (4.1)$$

В области  $D$  функцию  $p(x) = f(x) / \int_D f(x) dx$  можно рассматривать как плотность распределения вероятностей и тогда искомая точка максимума  $x^*$  является модой распределения  $p(x)$  [36,15]. Таким образом, эффективное моделирование плотности может существенно облегчить решение задачи (4.1).

Рассмотрим метод Метрополиса моделирования случайных величин распределенных с заданной плотностью вероятностей  $p(x)$ . В этом случае осуществляется моделирование цепи Маркова с подходящей переходной плотностью  $q(x, y)$ .

### 4.1. Выборка с отклонением

На практике, как правило,  $p(x)$  моделировать проблематично, поэтому имеет смысл воспользоваться для моделирования  $p(x)$  выборку с отклонением (rejection sampling). Для этого требуется подобрать *удобно моделируемую* плотность  $q(x)$ , для которой  $Cq(x) > f(x)$  для всех  $x \in D$ . Здесь эффективность метода будет зависеть от выбора константы  $C$  (выбор константы дает возможность не вычислять константу нормировки для  $f(x)$ ).

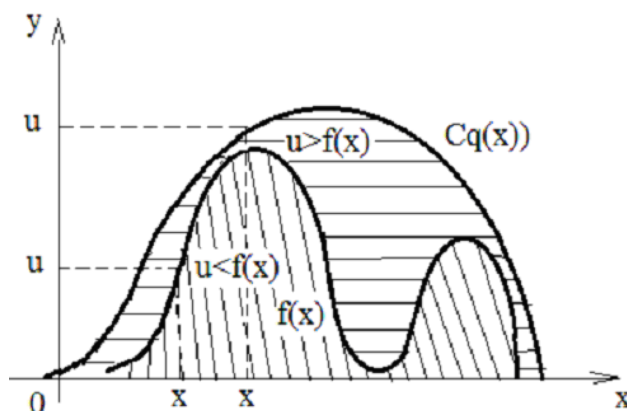


Рис. 4.1. Выборка с отклонением

Предположим, что константа  $C$  установлена. Алгоритм выборки с отклонением следующий [9,15].

**Алгоритм 1.**

1. Полагаем  $k = 1$ .
2. Моделируем плотность распределения  $q(x)$ , получаем  $X$ .
3. Моделируем случайную величину  $u$ , равномерно распределенную в промежутке  $[0, Cq(X)]$ .
4. Вычисляем  $f(X)$ .
  - А). Если  $f(X) > u$ , то  $X$  (запоминаем).
  - Б). Если  $f(X) \leq u$ , то  $X$  отклоняем.
5. Полагаем  $k = k + 1$  и переходим в пункт 2 пока  $k < N$  ( $N$  – некоторое заданное число итераций).

В результате Алгоритма 1 получим последовательность векторов

$$x^{(1)}, x^{(2)}, \dots, x^{(N)}, \tag{4.2}$$

являющихся реализациями  $p(x)$ , которые будут чаще попадать в окрестность  $x^*$ . По последовательности (4.2) несложно найти приближение  $x^* \approx x^{(k_0)}$ , где

$$f(x^{(k_0)}) = \max_{x^{(k)} \in D, k=1, N} f(x).$$

**4.2. Модифицированная выборка с отклонением**

В алгоритме выборки с отклонением возникает проблема с выбором константы  $C$ , которая заключается в следующем. Если  $C$  слишком большая константа, тогда алгоритм будет давать правильные результаты, но эффективность таких вычислений будет низкой. Если произошла ошибка в меньшую сторону, то область определения  $f(x)$  будет разбита на две части:

$$D = D_1 \cup D_2, \quad D_1 = \{x: f(x) \leq Cq(x)\}, \quad D_2 = \{x: f(x) > Cq(x)\}.$$

В области  $D_2$  выборка с отклонением применяться не может.

Чтобы устранить проблему с выбором константы  $C$ , можно модифицировать алгоритм выборки с отклонением [9].

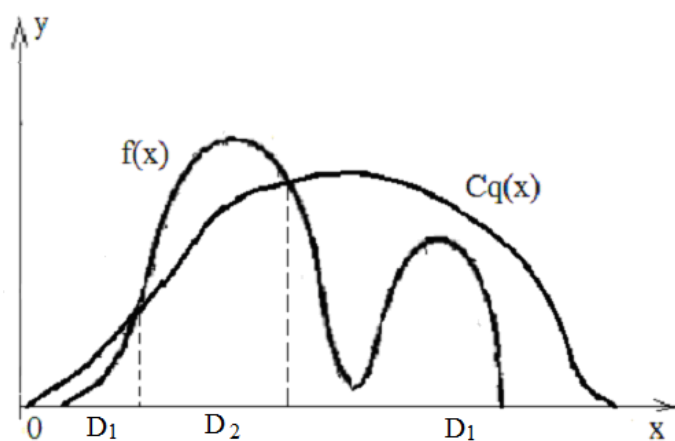


Рис. 4.2. Неудачный подбор константы  $C$

### Алгоритм 2.

1. Пусть  $k = 0$ , выбираем  $C_0$ .
2. Получаем реализацию  $X$ , моделируя  $q(x)$ .
3. Если  $X \in D_2$  (константу нужно уменьшить), то производим коррекцию константы:  $C_k = \frac{q(X)}{f(X)}$  и переходим к пункту 2; если  $X \in D_1$ , то проверяем неравенство

$$u \leq f(X), \quad (4.3)$$

где величина  $u$  равномерно распределена в промежутке  $[0, C_k q(X)]$ .

4. Если неравенство (4.3) выполнено, то  $x^{(k)} = X$ .
5. Переходим к пункту 2, не изменяя константы, и полагаем  $k = k + 1$ .

В этом алгоритме имеется также недостаток. Форма функции  $q(x)$  остается неизменной. Алгоритм Метрополиса исправляет это положение.

### 4.3. Метод Метрополиса

Алгоритм Метрополиса использует в качестве функции  $q$  переходную плотность марковского процесса, предполагая функцию  $q$ , зависящую от двух переменных  $q(x, y)$ . Вместо независимых реализаций случайной величины с плотностью  $q$  осуществляется моделирование цепи Маркова с переходной плотностью  $q(x, y)$ . Таким образом, алгоритм Метрополиса служит основой метода Монте-Карло на Марковских цепях.

### Алгоритм 3

Запишем алгоритм Метрополиса [34, 9,15].

1. Пусть  $k = 0$ . Выбираем произвольно  $X^{(0)}$  и переходную плотность  $q(x, y)$ . Моделируя  $q(X^{(0)}, y)$ , получаем  $Y^{(k)}$  ( $k = 0, 1, 2, \dots$ ).
2. Моделируя  $q(Y^{(k)}, y)$ , получим реализацию  $Y'$ .
3. Вычислим величину  $a = \frac{f(Y')}{f(Y^{(k)})} \frac{q(Y^{(k)}, X')}{q(Y', Y^{(k)})}$ .
4. Если  $a \geq 1$ , то полагаем  $Y^{(k+1)} = Y'$ .
5. Если  $a < 1$ , то с вероятностью  $a$  полагаем  $Y^{(k+1)} = Y'$  и с вероятностью  $a - 1$  полагаем  $Y^{(k+1)} = Y^{(k)}$ .

Случайный процесс  $Y^{(t)}$  с дискретным временем  $t = 0, 1, 2, 3, \dots$  является Марковским. Если цепь Маркова с переходной плотностью  $q(x, y)$  для некоторой плотности  $p(x)$  удовлетворяет соотношению

$$q(x, y)p(y) = q(y, x)p(x)$$

для всех  $x, y$ , то  $p(x)$  является плотностью стационарного распределения для  $q(x, y)$  [15]. В данном случае плотность  $p(x)$  пропорциональна функции  $f(x)$  и моделируя цепь Маркова, получаем реализации случайной величины с плотностью, пропорциональной функции  $f(x)$ . Поэтому метод Метрополиса может быть особенно эффективным при моделировании многомерных плотностей, для которых трудно подобрать мажоранту.

#### 4.4. Частный случай алгоритма Метрополиса

В качестве переходной плотности удобно выбрать гауссиан

$$q(x, y) = \frac{(\det B)^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(y-x)^T B^{-1}(y-x)},$$

для которого  $q(x, y) = q(y, x)$ . В этом случае  $a = \frac{f(Y')}{f(Y^{(k)})}$ .

Для простоты рассмотрим  $n = 1$  и тогда

$$q(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}}. \quad (4.4)$$

Условие  $a = \frac{f(Y')}{f(Y^{(k)})} \geq 1$  означает, что  $f(Y') > f(Y^{(k)})$  и по пункту 3 алгоритма 3 полагаем  $Y^{(k+1)} = Y'$  (будем продвигаться в сторону возрастания  $f(x)$ ).

Если  $a = \frac{f(Y')}{f(Y^{(k)})} < 1$  получаем, что  $f(Y') < f(Y^{(k)})$ . Возрастания нет. В этом случае с вероятностью  $a$  полагаем  $Y^{(k+1)} = Y'$  и с вероятностью  $1-a$  полагаем  $Y^{(k+1)} = Y^{(k)}$ . Для этого возьмем  $\eta_k$  – реализацию случайной величины, р.р. в  $(0,1)$ , и сделаем проверку. Если  $\eta_k \leq a$ , выполняем  $Y^{(k+1)} = Y'$ , если же  $\eta_k > a$ , то  $Y^{(k+1)} = Y^{(k)}$ .

Далее моделируем  $q(Y^{(k+1)}, Y)$  и получаем новый вектор  $Y'$  и т.д.

*Подбирая величину  $\sigma$  можно с достаточной точность продвинуться к моде  $p(x)$ .*

*Замечание.* В этом методе величины  $\sigma$ ,  $X_0$  являются произвольно назначенными величинами.

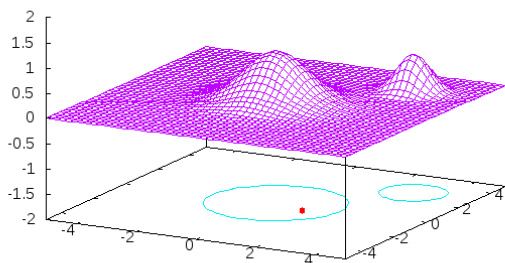


Рис.4.3 Метод Метрополиса. Первый шаг

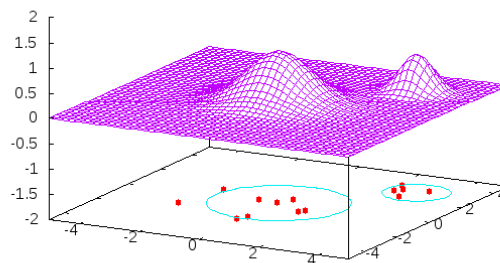


Рис.4.4 Метод Метрополиса. Последний шаг

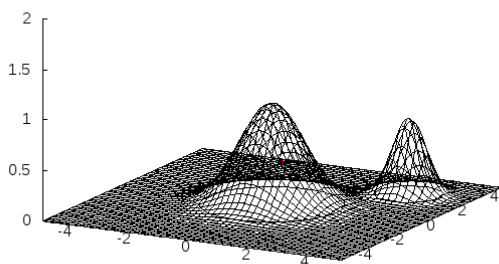


Рис.4.5 Метод Метрополиса. Первый шаг

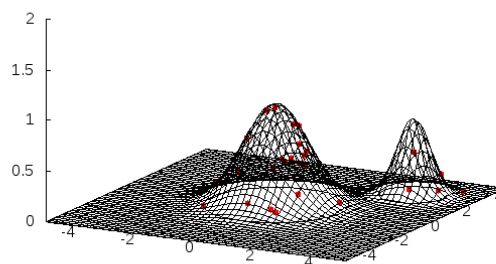


Рис.4.6 Метод Метрополиса. Последний шаг

## Глава V. Численные результаты

Будем численно решать задачу (1.10) нахождения оптимального плана. Как описано в главе II, эта задача сводится к нахождению максимума функции  $A(x_1, x_2, x_3, x_4)$  в области точек плана  $x_1, x_2, x_3, x_4 \in [0, B]$ . Применим для нахождения глобального экстремума случайный поиск с “памятью”, описанный в главе III. Он использует моделирование нормально распределенных случайных векторов. В результате пробные точки в основном будут сосредоточены в эллипсоиде рассеяния, центр которого является приближением экстремальной точки, полученной на предыдущей итерации. На нулевой итерации применяется простейший случайный поиск. Матрица ковариаций, элементы которой оцениваются в процессе поиска экстремума, несет в себе информацию о целевой функции. В связи с этим метод имеет название с “памятью”. Это самообучающийся метод, так как осуществляет уточнение области поиска от итерации к итерации.

**Таблица 5.1**

	1	2	3	4	5	6
$i$	$\Delta_i$	$x_1^{(i)}$	$x_2^{(i)}$	$x_3^{(i)}$	$x_4^{(i)}$	$A(x^{(i)}, \Delta_i)$
1	0.1	0.078371	0.568601	1.87264	4.28987	4.20E-11
2	0.2	0.02889	0.477314	1.2182	3.56277	1.28E-08
3	0.3	0.087908	0.500682	2.02852	3.61996	2.47E-07
4	0.4	0.038685	0.577232	1.80619	3.7572	5.14E-06
5	0.5	0.013134	0.562389	1.37124	4.17177	4.19E-05
6	0.6	0.031447	0.649374	2.00118	5.37435	0.000271717
7	0.7	0.001671	0.661616	2.79935	7.27505	0.00135662
8	0.8	0.034846	0.473323	2.10385	9.65124	0.0089874
9	0.9	0.040175	0.463833	2.18561	8.40895	0.0834351

Метод случайного поиска глобального экстремума с “памятью” хорошо зарекомендовал

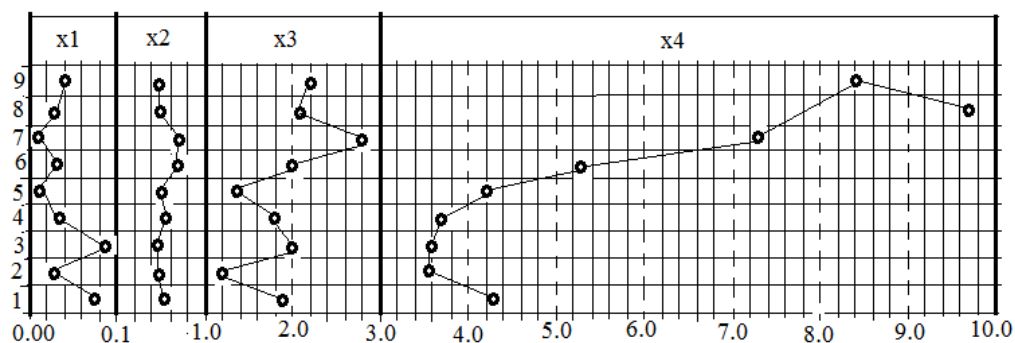


Рис.5.1. Оптимальные планы, полученные при использовании случайного поиска глобального экстремума с “памятью” для 9 вариантов смеси

себя [13,14], так как он позволяет находить глобальный экстремум за 2-3 итерации для многомерных функций с большим числом локальных экстремумов. В [36] приводится теорема о сходимости этого метода.

Численные результаты работы на ЭВМ программы с алгоритмом метода с “памятью” помещены в таблицу 5.1. Для каждого из различных вариантов истинных значений нелинейных параметров по (1.11) получаем свой оптимальный план. Рассмотрено 9 вариантов.

Если в реальной задаче указаны априори значения нелинейных параметров, то можно подобрать подходящий план.

В целом, из рис. 5.1 можно сделать следующий вывод. Для выявления из смеси одного изотопа следует проводить измерения в следующие промежутки времени

$$\begin{aligned} x_1 &\in [0.01; 0.09], \quad \Delta x_1 = 0.08, \\ x_2 &\in [0.14; 0.17], \quad \Delta x_2 = 0.03, \\ x_3 &\in [1.2; 2.8], \quad \Delta x_3 = 1.6, \\ x_4 &\in [3.6; 9.7], \quad \Delta x_4 = 6.1, \end{aligned}$$

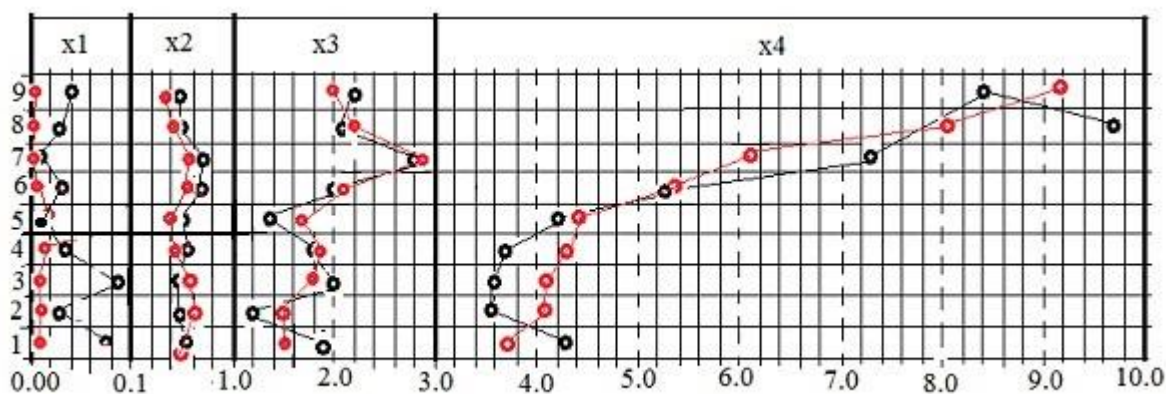
где  $\Delta x_i$  – длина промежутка для  $i$ -го измерения  $i = \overline{1,4}$ . Получается следующая закономерность: первое измерение нужно проводить в очень узкий промежуток времени от начала процесса, второй – еще в более узкий промежуток времени, для третьего – промежуток измерений увеличивается, а для четвертого измерения момент измерения увеличивается с увеличением разности между нелинейными параметрами и приближается к моменту времени, равному двум максимальным (из двух) полупериодам полураспада.

Теперь приведем результаты решения задачи (1.10) нахождения оптимального плана, с использованием метода Метрополиса, который представлен в главе IV. Также, как и при оптимизации на основе случайного поиска глобального экстремума с “памятью”, рассмотрим 9 вариантов значений нелинейных параметров по правилу (1.11). Результаты счета приведены в таблице 5.2.

**Таблица 5.2**

	1	2	3	4	5	6
$i$	$\Delta_i$	$x_1^{(i)}$	$x_2^{(i)}$	$x_3^{(i)}$	$x_4^{(i)}$	$A(x^{(i)}, \Delta_i)$
1	0.1	0.0070	0.4847	1.5270	3.7391	7.4934E-11
2	0.2	0.0059	0.5936	1.5879	4.0824	1.9426E-08
3	0.3	0.0067	0.5557	1.8193	4.1242	5.9973E-07
4	0.4	0.0134	0.4044	1.8242	4.3500	6.7032E-06
5	0.5	0.0194	0.4334	1.6866	4.2177	5.0583E-05
6	0.6	0.0024	0.6130	2.0849	5.3092	3.4001E-04
7	0.7	0.0014	0.5838	1.8836	6.1730	2.0410E-03
8	0.8	0.0058	0.5167	2.1630	8.0973	1.3077E-02
9	0.9	0.0042	0.3888	2.0392	9.2940	1.1829E-01

Рис. 5.2 наглядно демонстрирует, что оптимальные планы, полученные двумя способами оптимизирования функции (1.10), отличаются, но характер измерения для каждого  $x_i$  сохраняется.



**Рис. 5.2. Сравнение оптимальных планов, полученных при использовании**

На рис. 5.2 (красные точки) во всех вариантах по оптимальному плану, полученному при использовании метода Метрополиса, первое измерение следует проводить сразу же в начале процесса. Второе измерение имеет очень маленький разброс по времени и можно сказать, что для всех вариантов его можно производить при  $x_2^i \approx 1.5$ ,  $i = \overline{1,9}$ .



## Выводы

В настоящей работе рассматривается нелинейная экспоненциальная регрессия и ставится задача получения локально-оптимального плана.

1. В главе II получена информационная матрица и через эту матрицу получен  $D$ -оптимальный критерий, который представляет собой функцию, зависящую от нелинейных параметров регрессии и от точек плана.
2. Для оптимизации критерия использовались два метода случайного поиска глобального экстремума.

А). Первый метод описан в главе III. Это итерационный метод. Он называется случайный метод с “памятью”. Этот метод основан на моделировании случайного нормально распределенного вектора и оценивании элементов матрицы ковариаций, с помощью которой от итерации к итерации область поиска сужается и в конечном итоге превращается в окрестность приближенно экстремальной точки, которая уточняется от итерации к итерации. Рассмотрены две тестируемые функции. Рис. 3.1-3.10 демонстрируют этот процесс. Метод сходится с увеличением пробных точек.

Б). Второй метод представлен в главе IV. Это метод Метрополиса. При использовании этого метода требуется моделировать цепи Маркова и по определенному алгоритму продвигаться к экстремальной точке. В качестве переходной плотности берется Гауссиан. В этом случае алгоритм упрощается.

На рис. 4.3-4.6 на тестовой функции, являющейся суммой двух Гауссианов, представлены варианты поиска экстремума. При большом числе итераций плотностью стационарного распределения для переходной плотности имеет плотность распределения пропорциональную функции  $f(x)$  и моделируя цепь Маркова, получаем реализации случайной величины с плотностью, пропорциональной функции  $f(x)$ . Поэтому метод Метрополиса оказался эффективным.

Второй метод показался проще для реализации, однако, первый метод легко поддается распараллеливанию (эксперименты проведены в бакалаврской работе для линейной регрессии).

3. Для 9 вариантов нелинейных параметров получены соответствующие локально-оптимальные планы. Результаты представлены в главе V. На рис. 5.2 легко увидеть промежутки, в которые следует проводить измерения. Они занимают определенный интервал времени для каждого измерения. Особенность в промежутке для последнего измерения заключается в том, что чем больше разница между полупериодами распада радионуклидов, тем дальше время последнего измерения.

## Заключение

В диссертационной работе для нелинейной экспоненциальной регрессии рассмотрена задача построения локально оптимальных планов. Для этого выбран определенный критерий оптимальности – это критерий  $D$  - оптимальности. Он основан на предположении, что погрешности между результатами экспериментов в точках множества планирования и значениями регрессионной функции в этих же точках множества планирования есть случайные величины, распределенные нормально.

В этом случае при достаточно большом числе экспериментов оценка неизвестных параметров нелинейной регрессии по методу наименьших квадратов стремится к истинным параметрам, а дисперсия вектора погрешностей будет определяться через информационную матрицу. Поэтому объем эллипсоида рассеяния вектора погрешностей тоже определяется через информационную матрицу. Уменьшая этот объем, будем уменьшать и величины погрешностей для оцененных параметров. Отсюда появляется функция, которую требуется оптимизировать.

Для нелинейной регрессии эта функция зависит не только от точек плана, но и от самих параметров. Поэтому задача оптимизации решается при некоторых параметрах регрессии, которые априори известны. Полученный план называется локально оптимальным.

Для оптимизации использовались два случайных метода поиска глобального экстремума. Оба метода позволяют находить реализации случайных векторов, распределенных с плотностью пропорциональной целевой функции. Поэтому задачу нахождения глобального максимума можно трактовать как нахождение моды плотности.

Для апробации метода нахождения локально оптимальных планов рассматривалась задача определения радионуклидов в смеси из двух. Рассмотрено 9 вариантов смесей радионуклидов.

Результат получен двумя способами оптимизации и наглядно представлен рисунком. Легко увидеть промежутки, в которые следует проводить измерения. Они занимают определенный промежуток для каждого измерения. Особенность в промежутке для последнего измерения заключается в том, что чем больше разница между полупериодами распада радионуклидов, тем дальше время последнего измерения.

Таким образом, применение метода планирования эксперимента может дать значительный экономический эффект, так как получен соответствующий локально оптимальный план.

## Список литературы

1. Бекман И.Н. Ядерная физика. М., Изд-во МГУ, 2005, 867 с.
2. Буре В.М., Кирпичников Б.К., Конспект-справочник лекций по теории вероятностей и математической статистике, СПб, 1999, с. 143.
3. Буре В.М., Ковригин А.Б., Седунов Е.Б. Критерии минимаксного типа в несмещенном планировании регрессионных экспериментов. Вопросы кибернетики, 1981, № 73. С. 69-83.
4. Бусленко Н.П., Шрейдер Ю.А. Метод статистических испытаний. Физматгиз, 1961.
5. Владимирова Л.В. Использование случайного поиска с “памятью” в оценке параметров нелинейной регрессии. Вестник СПГУТД, №4, 2013, с. 30-34.
6. Владимирова Л.В. Методы Монте-Карло в задаче оптимизации динамики пучков. Вестник Санкт-Петербургского университета, сер.10, 2014.
7. Владимирова Л.В., Фатьянова И.А Построение оптимального плана регрессионного эксперимента на основе случайного поиска с “памятью” с использованием параллельных вычислений. С. 303-304. Устойчивость и процессы управления: Материалы III международной конференции (Санкт-Петербург, 5-9 октября 2015 г.)/Под ред. А.П. Жабко, Л.А. Петросяна. СПб.: Издательский Дом Федоровой Г.В., 2015. – 623 с.
8. Войтишек А.В., Михайлов Г.А. Численное статистическое моделирование. Методы Монте-Карло. М. “Академия”, 2006. 368 с.
9. Ермаков С.М. Метод Монте-Карло в вычислительной математике: Вводный курс.–СПб.:Невский диалект; М. : БИНОМ. Лаборатория знаний, 2009. – 192 с.
10. Ермаков С.М. Математическая теория планирования эксперимента. Изд-во “Наука”, М. 1983, 392 с.
11. Ермаков С.М. Методы Монте-Карло и смежные вопросы. Изд-во

“Наука”, М. 1975, 472 с.

12. Ермаков С.М., Жиглявский А.А. Математическая теория оптимального эксперимента. М., “Наука”, 1987. 320 с.

13. Ермаков С.М., Владимиров Л.В. Об одном методе поиска экстремума, основанном на оценивании ковариационной матрицы. Автоматика и вычислительная техника, №5, Рига, 1977. С. 38-41.

14. Ермаков С.М., Митиоглова Л.В. Об одном методе поиска экстремума, использующем обобщенный метод Неймана моделирования случайных величин. Методы Монте-Карло в вычислительной математике и математической физике (Сб.), Новосибирск, 1974, 146-151.

15. Ермаков С.М., Сипин А.С. Метод Монте-Карло и параметрическая разделимость алгоритмов. Изд-во СПбГУ, 2014. 248 с.

16. Жиглявский А.А. Математическая теория глобального случайного поиска. Изд-во ЛГУ, 1985, 296 с.

17. Козлов В.П. Избранные труды по теории планирования эксперимента и обратным задачам оптического зондирования. СПб, 2000. 500 с.

18. Колмогоров А.Н. Теория передачи информации. М., изд. АН СССР, 1956.

19. Линник Ю.В. Метод наименьших квадратов и основы теории обработки наблюдений, Физматгиз, М., 1962.

20. Математическая теория планирования эксперимента. СМБ под редакцией С.М.Ермакова. М. “Наука”, 1983. 391 с.

21. Мелас В.Б. Одна теорема двойственности и E-оптимальность. Заводская лаборатория, 1982, № 3, с. 48-50.

22. Мелас В.Б. Оптимальное планирование эксперимента для экспоненциальной регрессии// Математические методы планирования эксперимента. Новосибирск: Наука, 1981. С. 174-198.

23. Растринин Л.А. Статистические методы поиска. М., Наука, 1968.

24. Седунов Е.В. Планирование эксперимента: неадекватные модели. С-Петербург, 2009. 208 с.

25. Сетина (Фатьянова) И.А., Владимирова Л.В. Поиск локально оптимальных планов для экспоненциальной регрессии. 48 международная научная конференция аспирантов и студентов “Процессы управления и устойчивость”, 3-6 апреля 2017 г.
26. Толстой Н.А., Теофилов П.П. Новый метод исследования релаксационных процессов и его применение к изучению некоторых физических явлений. – Успехи физ. Наук, 1950, т.41, с. 44-107.
27. Федоров В.В. Теория оптимального эксперимента. М., “Наука”, 1971. 312 с.
28. Цыбаков А.Б. Рекуррентное оценивание моды многомерного распределения. Проблемы передачи информации. 1990. Т.26. Вып. 1. С. 38-45.
29. Box G. A basis for the selection of a response surface design. J. Amer. Statist. Assos. 1959. V. 54. P. 662-654.
30. Chernoff H. Locally optimal design for estimator parameters// Ann. Math.Stat. 1953. Vol. 24. P.586-602.
31. Kiefer J. General equivalence theory for optimum design (approximate theory). Annals Statist. 1974, №5. P. 849-879.
32. Fisher R. The design of experiments. London, 1960. 248 p.
33. Melas V.B. Optimal designs for exponential regression// Math. Operations forsh. Statist. Ser. Statistics. 1978. Vol. 9. P. 45-59.
34. Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E. Equations of state calculations by fast computing machines // Journal Chemical Physics. Vol. 21. 1953. P. 1087-1091.
35. Seber G.A.F., Wild C.G. Nonlinear Regression. New York: Wiley, 1989.
36. Vladimirova L.V. Global Extremum Search on the Basis of Density and Its Mode Estimation. Proc. of 20<sup>th</sup> International Workshop: Beam Dynamics & Optimization (BDO), June-July 2014, St. Petersburg, Russia. IEEE, 2014, p. 186.
37. Vladimirova L., Fatyanova I. Construction of Regression Experiment Op-

timal Plan Using Parallel Computing Proceedings of III International Conference “Stability and Control Processes” in Memory of V.I.Zubov (SCP), IEEE, 2015. Pp. 361-363. DOI: 10.1109/SCP.2015.7342140

38. Wiese W.L. Atomic transition probabilities. – Nucl. Inst. And Methods, 1970, v. 90, p. 25.

## Приложения

### Приложение 1. Моделирование равномерно распределенного вектора в заданной области

Определим параллелепипед  $\Pi$ , в котором содержалась бы область  $D$ :

$D \subset \Pi$ ,  $\Pi = \{a_i \leq x_i \leq b_i, i = \overline{1, n}\}$ . Точка  $\xi = (\xi_1, \dots, \xi_n)$ , равномерно распределенная в параллелепипеде  $\Pi$ , моделируется по формулам

$$\xi_i = (b_i - a_i)\gamma_i + a_i, \quad i = \overline{1, n}, \quad (\text{П.1.1})$$

где  $\gamma_i, i = \overline{1, n}$  – независимые реализации равномерно распределенной в  $(0, 1)$  случайной величины  $\gamma$ . Если полученная точка  $\xi = (\xi_1, \dots, \xi_n)$  принадлежит области  $D$ , то мы имеем реализацию равномерно распределенного случайного вектора в этой области. В противном случае, когда полученная реализация случайного вектора не принадлежит  $D$ , получаем другую реализацию  $\xi = (\xi_1, \dots, \xi_n)$  по (2.2), используя новые значения  $\gamma_i, i = \overline{1, n}$ . Эта процедура повторяется до тех пор, пока не получим  $N$  реализаций случайного равномерно распределенного вектора в области  $D$ :

$$\xi^{(i)} = (\xi_1^{(i)}, \dots, \xi_n^{(i)}), \quad i = \overline{1, N}. \quad (\text{П.1.2})$$

### Приложение 2. Метод Монте-Карло для оценки интегралов

Метод Монте-Карло является по существу методом приближенного интегрирования [11]. Пусть случайная величина  $X$ , принимающая значения в некоторой заданной области  $G \subset E^n$ , имеет плотность распределения вероятностей  $\rho(x)$ . Введём некоторую случайную величину  $\eta = \Phi(X)$ . Ее математическое ожидание равно

$$M\eta = \int_G \Phi(x)\rho(x)dx.$$

По закону больших чисел с вероятностью 1 имеем

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \Phi(x^{(k)}) = \int_G \Phi(x)\rho(x)dx, \quad (\text{П.2.1})$$

где  $x^{(k)}, k = \overline{1, N}$  – независимые реализации случайной величины  $X$ , распределённой в области  $G$  с плотностью вероятностей  $\rho(x)$ ;  $N$  – число реализаций  $x^{(k)}$ .

Пусть существует конечный второй момент случайной величины  $\eta = \Phi(X)$ . Тогда в качестве частного случая центральной предельной теоремы при тех же предположениях относительно  $x^{(k)}$ , что и выше, имеем

$$\lim_{N \rightarrow \infty} P\left(\frac{\sqrt{N}}{\sigma} \left| \frac{1}{N} \sum_{k=1}^N \Phi(x^{(k)}) - \int_G \Phi(x)\rho(x)dx \right| < z\right) = \sqrt{\frac{2}{\pi}} \int_0^z e^{-u^2} du,$$

где  $\sigma^2 = \int_G \Phi^2(x)\rho(x)dx - \left( \int_G \Phi(x)\rho(x)dx \right)^2$  – дисперсия случайной величины  $\eta = \Phi(X)$ , или с заданной вероятностью  $\gamma$  неравенство

$$\left| \frac{1}{N} \sum_{k=1}^N \Phi(x^{(k)}) - \int_G \Phi(x)\rho(x)dx \right| \leq \frac{z\sigma}{\sqrt{N}}. \quad (\text{П.2.2})$$

Значение  $z$  определяется по заданной доверительной вероятности  $\gamma$

$$\int_0^z e^{-u^2} du = \sqrt{\frac{\pi}{2}}\gamma.$$

Следовательно, если интеграл  $\int_G \Phi(x)\rho(x)dx$  оценивается с помощью среднего арифметического

$\frac{1}{N} \sum_{k=1}^N \Phi(x^{(k)})$ , то погрешность при заданной вероятности  $\gamma$  убывает как  $O\left(\frac{1}{\sqrt{N}}\right)$ .

Формулы (П.2.1), (П.2.2) служат обоснованием использования суммы

$\frac{1}{N} \sum_{k=1}^N \Phi(x^{(k)})$  в качестве приближённого значения интеграла.

В результате имеем формулу для приближенного вычисления интеграла

$$\int_G \Phi(x)\rho(x)dx = \frac{1}{N} \sum_{k=1}^N \Phi(x^{(k)}) \quad (\text{П.2.3})$$



### **Приложение 3. Моделирование нормально распределенного случайного вектора [11]**

#### **1). Моделирование двумерного нормального распределения**

Рассмотрим две независимые нормально распределённые случайные величины  $\xi, \eta$  с параметрами  $a = 0, \sigma = 1$ . Следовательно, их плотности распределения соответственно равны

$$p_{\xi}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad p_{\eta}(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}},$$

а совместная плотность распределения этих величин, так как они независимы, есть функция

$$p(x, y) = p_{\xi}(x) \cdot p_{\eta}(y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}.$$

Перейдём к полярной системе координат

$$\begin{aligned} x &= r \cos \varphi, \\ y &= r \sin \varphi, \end{aligned} \tag{П.3.1}$$

причём  $r \in [0, \infty)$ ,  $\varphi \in [0, 2\pi)$ .

Тогда в полярной системе координат плотность распределения случайной точки будет иметь вид

$$\tilde{p}(r, \varphi) = \frac{1}{2\pi} r e^{-\frac{r^2}{2}}.$$

Представим эту плотность в виде

$$\tilde{p}(r, \varphi) = \frac{1}{2\pi} \cdot r e^{-\frac{r^2}{2}} = p_{\varphi}(\varphi) \cdot p_r(r).$$

То есть, плотность в полярной системе координат также представляется в виде произведения двух плотностей и, следовательно, распределения полярных координат являются независимыми. Далее получим формулы моделирования для переменных  $r$  и  $\varphi$  из соотношений:

$$\int_0^r t e^{-\frac{t^2}{2}} dt = \gamma_1,$$

$$\int_0^{\varphi} \frac{1}{2\pi} dt = \gamma_2.$$

Здесь  $\gamma_1, \gamma_2$  независимые реализации случайной равномерно распределенной в  $(0,1)$  величины  $\gamma$ . Решая эти уравнения относительно  $r$  и  $\varphi$ , получаем следующие формулы моделирования

$$r = \sqrt{-2 \ln \gamma_1},$$

$$\varphi = 2\pi\gamma_2. \quad (\text{П.3.2})$$

Подставим П.3.2 в П.3.1 и получим искомые формулы

$$\xi = \sqrt{-2 \ln \gamma_1} \cos(2\pi\gamma_2),$$

$$\eta = \sqrt{-2 \ln \gamma_1} \sin(2\pi\gamma_2). \quad (\text{П.3.3})$$

моделирования случайного вектора  $(\xi, \eta)$ , нормально распределённого на плоскости.

## 2). Моделирование многомерного нормального распределения

Известно, что плотность вероятностей нормально распределённого случайного вектора  $\xi = (\xi_1, \dots, \xi_n)$  с параметрами  $(a, C)$  равна

$$p(x) = \frac{(\det C)^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}(x-a)^T C^{-1}(x-a)\right). \quad (\text{П.3.4})$$

Здесь  $a = M\xi = (a_1, a_2, \dots, a_n)$  – математическое ожидание случайного вектора  $\xi = (\xi_1, \dots, \xi_n)$ , а  $C = \{c_{ij}\}_{i,j=1}^n$  – положительно определённая и симметрическая матрица, матрица ковариаций:

$$a_i = \int_D x_i p(x) dx,$$

$$c_{i,j} = \int_D x_i x_j f(x) dx - \int_D x_i f(x) dx \int_D x_j f(x) dx, \quad i, j = \overline{1, n}. \quad (\text{П.3.5})$$

Пусть для простоты изложения вектор средних равен нулю:  $a = 0$ .

Матрицу ковариаций можно представить через произведение матрицы  $S$ , строки которой являются собственными векторами, и матрицы  $D$  – диагональной матрицы, по диагонали которой располагаются собственные числа:

$$C = SDS^T = SD^{\frac{1}{2}} D^{\frac{1}{2}} S^T = SD^{\frac{1}{2}} \left( SD^{\frac{1}{2}} \right)^T. \quad (\text{П.3.6})$$

Преобразуем аргумент экспоненты в П.3.4, используя П.3.6

$$x^T \left[ SD^{\frac{1}{2}} \left( SD^{\frac{1}{2}} \right)^T \right]^{-1} x = x^T \left[ \left( SD^{\frac{1}{2}} \right)^T \right]^{-1} \left[ SD^{\frac{1}{2}} \right]^{-1} x =$$

$$= \left[ \left( SD^{\frac{1}{2}} \right)^{-1} x \right]^T \left[ SD^{\frac{1}{2}} \right]^{-1} x = y^T E y,$$

где

$$y = \left( SD^{\frac{1}{2}} \right)^{-1} x. \quad (\text{П.3.7})$$

Откуда имеем

$$x = SD^{\frac{1}{2}} y. \quad (\text{П.3.8})$$

Формула П.3.8 является формулой моделирования для нормально распределённого вектора с параметрами  $(0, C)$ .

Вектор

$$x = SD^{\frac{1}{2}} y + a \quad (\text{П.3.9})$$

нормально распределен со средним  $a$  и матрицей ковариаций  $C$ .

**Замечание.**

Компоненты вектора  $y = (y_1, \dots, y_n)$  можно получать и по формулам моделирования П.3.3 (формула моделирования случайного нормально распределенного двумерного вектора). Обращаясь один раз к П.3.3 получим две компоненты  $y$ . Если  $n > 2$ , то нужно несколько раз использовать формулы П.3.3, пока не получим все компоненты вектора  $y$ .

## Приложение 4. Листинг программы.

```
#include <QCoreApplication>
#include <cmath>
#include <cstdlib>
#include <cstdio>
#include <ctime>
#include <fstream>
#include <iostream>
#include <sstream>
#include <string>
#include <vector>
#include <gsl/gsl_math.h>
#include <gsl/gsl_rng.h>
#include <gsl/gsl_eigen.h>
#include <gsl/gsl_blas.h>
#include <gsl/gsl_matrix.h>
#include <gsl/gsl_linalg.h>
#include <gsl/gsl_randist.h>
#include <sys/time.h>
using namespace std;
#include "random.h"
using namespace cpt;
const string title("Metropolis Algorithm for Two Gaussians");
Random rg; // random number generator
double Theta1=1.8 ,Theta2=1.2;
long mtime()
{
    struct timeval t;

    gettimeofday(&t, NULL);
    long mt = (long)t.tv_sec * 1000 + t.tv_usec / 1000;
    return mt;
}
double F(double*y)
{
    gsl_matrix *Fmatr=gsl_matrix_alloc(4,4);
    double*x=new double [4];
    for (int i=0;i<4;i++) x[i]=y[i];
    x[0]=0;
    for(int i=0;i<4;i++)
    {
        gsl_matrix_set(Fmatr,0,i,exp(-Theta1*x[i]));
        gsl_matrix_set(Fmatr,1,i,-x[i]*exp(-Theta1*x[i]));
        gsl_matrix_set(Fmatr,2,i,exp(-Theta2*x[i]));
        gsl_matrix_set(Fmatr,3,i,-x[i]*exp(-Theta2*x[i]));
    }
    int s(0);
    gsl_permutation*p=gsl_permutation_alloc(4);
    gsl_linalg_LU_decomp(Fmatr,p,&s);
    double f=pow(gsl_linalg_LU_det(Fmatr,s),2);
    gsl_permutation_free(p);
    gsl_matrix_free(Fmatr);
    return f;
}
int MaxFunc(double F,double*X,double *MaxF,double*MaxX,int n)
{
    double mf=*MaxF;
    if(F>mf)
    {
        *MaxF=F;
        for(int i=0;i<n;i++)
            MaxX[i]=X[i];
    }
    return 0;
}
double
    MaxX[4] = {0.0, 0.0, 0.0, 0.0},
    MaxF=0;
// Metropolis
Point Walker0{1,1,1,1}; // initial position
double Walker [4] = {1,1,1,1};
double delta = 1.0;
int stepsToDiscard = 0;
unsigned long steps = 0;
unsigned long accepts = 0; // steps accepted
Gnuplot gnuplot; // initialize Gnuplot instance
```

```

int stepsToSkip = 200;          // when plotting

double P (double x, double y) {
    double p = 0;
    for (int i = 0; i < 2; i++)
        p += A[i]*exp(-((x-center[i])*(x-center[i])/(2*sigma[i]*sigma[i]) + (y-center[i])*(y-center[i])/(2*sigma[i]*sigma[i]))));
    return p;
}

double P (double *x) {
    double p = 0;
    p=F(x);
    return p;
}

bool MetropolisStep ( ) {      //is the step accepted?
    double Trial[4];
    for(int i=0;i<1000000000000000;i++){

        for (int i = 0; i < 4; i++)
            Trial[i] = rg.rand_gauss(Walker[i],5);

        if(Trial[0]>0.f && Trial[0]<10.f&&Trial[1]>0.f && Trial[1]<10.f&&Trial[2]>0.f && Trial[2]<10.f&&Trial[3]>0.f && Trial[3]<10.f)
            break;
    }

    double ratio = P(Trial)/P(Walker);

    if (rg.rand() < ratio) {

        MaxFunc(F(Trial),Trial,&MaxF,MaxX,4);
        for (int i = 0; i < 4; i++){
            Walker[i]= Trial[i];
            return true;
        } else
            return false;
    }
}

void MonteCarloStep ( ) {
    for (int i = 0; i < stepsToDiscard; i++)
        MetropolisStep();
    if (MetropolisStep())
        ++accepts;
    ++steps;
}

void takeStep (int stepno) {

    for (int i = 0; i <= stepsToSkip; i++) {
        MonteCarloStep();
    }

    ostringstream os;
    double percent = int(accepts / double(steps) * 100.0);
    os << "set title \" " << title
        << ", Metropolis steps = " << steps
        << ", accepts = " << percent << "%\';
    gnuplot.command(os.str());
}

int main () {
    cout << " " << title << endl;
    cout << " Enter starting position x_0, y_0: ";
    cin >> Walker[i]>>;
    cout << " Enter number of Monte Carlo steps: ";
    int n;
    cin >> n;
    n=100;
    stepsToDiscard=100;
    cout << " Enter starting position Theta1, Theta2: ";
    cin >> Theta1 >> Theta2;
    // set random number generator seed
    rg.set_seed time();
    cout << " Using " << rg.get_algorithm()
        << " and seed " << rg.get_seed() << endl;

    long t = mtime();
    for (int i = 0; i < n; i++)
        takeStep(i);
    t = mtime() - t;
    cout<<endl<<"All progr: "<<<double(t)/1000<<" sec."<<endl;
}

```

```
cout<<"Teor= "<<F(x2)<<endl;
cout<<"MAX_F="<<MaxF;
cout<<"_____MaxX=[";
for(int i = 0; i < 4; i++)
cout<<MaxX[i]<<" ";
cout<<"]"<<endl;
```

```
}
```