

В. Н. Троян

ОБРАТНЫЕ ГЕОФИЗИЧЕСКИЕ
ЗАДАЧИ

Курс лекций

Санкт-Петербургский университет
2001

Введение

Решение геофизических обратных задач предполагает совместное рассмотрение результатов измерений, теоретических соотношений между измеряемыми величинами и параметрами геофизических объектов (решение прямой задачи), а также априорной информации относительно данных и моделей. В геофизическом анализе очень важен выбор модели для решения прямой задачи. Такая модель должна адекватно описывать связь между измеряемыми величинами и параметрами геофизических объектов. Кроме этого очень важно определить общее число параметров и выделить те параметры, изменение которых наиболее сильно влияет на измеряемые величины. Естественно, что модель описывающая связь между измеряемыми величинами и параметрами геофизических объектов определяется типом решаемой задачи и геологическими (геофизическими) характеристиками района исследований. При заданных результатах наблюдений решение обратной задачи заключается в определении геологической (геофизической) модели, которая наилучшим образом согласуется с результатами наблюдений. Фактически обращение состоит в нахождении таких геофизических параметров, которые доставляют наилучшее согласие с экспериментальными данными. Следовательно, выбор модели является важным как для решения прямой задачи, так и для ее обращения. Первостепенное значение при этом имеет физическое обоснование решения прямой и обратной задачи. Экспериментальные данные обычно бывают искажены различного рода факторами не поддающимися учету. Кроме этого, всегда имеется расхождение между реальными физическими процессами и описывающими их моделями. Вероятностно-статистический подход к решению обратных задач как раз и предполагает существование расхождений между модельными и экспериментальными данными.

Необходимость решения обратных задач возникает в различных областях человеческой деятельности. Так, например, следователь должен восстановить картину преступления по ряду косвенных улик. Кардиолог - определить состояние сердца больного по кардиограмме. В последние десятилетия в медицине для восстановления образов различных частей тела человека применяются томографические методы. Томографические методы демонстрируют высокую эффективность и геофизике при изучении внутреннего строения Земли, используя геофизические данные, полученные на поверхности Земли и в скважинах.

Обратные задачи имеют много общих особенностей и близкие методы решений вне зависимости от физической природы рассматриваемых явлений:

- Обратные задачи являются некорректными, т.е. небольшие изменения во входных данных могут приводить к большим отклонениям в решении.
- Решение обратных задач не обладает свойством единственности, т.е. исходные данные не содержат достаточной информации для получения единственного решения.
- Самая важная особенность обратных задач – присутствие случайного шума

в реальных геофизических данных, что приводит к необходимости применения вероятностно статистических методов.

При решении обратных задач очень важно обратить внимание на следующее:

Построение модели измерений.

Роль априорной информации в решении обратной задачи.

Метод решения конкретной геофизической задачи.

Качество решения задачи.

В рамках предлагаемого курса мы хотели получить ответы на эти и многие другие проблемы.

Оставаясь на принципах статистического подхода к интерпретации геофизических данных мы дадим наиболее важные теоретические результаты решения обратных геофизических проблем и рассмотрим их практическое применение.

Краткое содержание предлагаемого курса:

Введение в теорию вероятностей. Определение вероятности, формула Байеса, распределение вероятностей (нормальное, равномерное, Лапласа и т.д.), числовые характеристики распределения (математическое ожидание, медиана, дисперсия, коэффициент корреляции, ковариационная матрица, доверительный интервал).

Информация и энтропия. Информация Шеннона и информация Фишера, матрица информации Фишера.

Случайные функции и их свойства. Случайные процессы и временные ряды, автокорреляционная и взаимокорреляционная функции, спектр мощности и корреляционная функция.

Элементы математической статистики. Выборка, оценка, робастное оценивание, неравенство Рао–Крамера.

Модели измерений. Пространство моделей для описания поля искомым параметров, например, скорости Р- и S-волн, плотности, удельной проводимости и т.д; построение модели данных измерений; примеры решения прямых и обратных задач; влияние ошибок в выборе модели и ошибок измерения; введение априорной информации относительно неизвестных параметров.

Решение обратных геофизических задач: линейные модели. Методы оценки параметров при решении линейных обратных задач: метод наименьших квадратов, метода максимального правдоподобия, сингулярный анализ, метод наименьших модулей, метод Байеса, методы регуляризации и учет априорной информации.

Решение обратных геофизических задач: нелинейные модели. Методы оценки параметров при решении нелинейных задач: метод Ньютона, метод Ньютона–Лекама, процедуры линеаризации, выбор вектора начальных параметров, генетический алгоритм.

Статистические критерии выбора модели. Проверка параметрических гипотез, критерий отношения апостериорных вероятностей, задача разделения сигналов.

Лучевая геофизическая томография. Лучевая томография: преобразование Радона, преобразование Гильберта, обратное проецирование, алгебраический метод, рекуррентная процедура, восстановление медленности и коэффициента поглощения.

Дифракционная томография. Метод возмущения для скалярного волнового уравнения, аппроксимация Борна, томографический функционал, алгоритм реконструкции.

Многофакторный анализ и обработка временных рядов. Анализ и выбор главных компонент.

Предложенный курс лекций иллюстрируется компьютерными примерами и компьютерными упражнениями.

Глава 1

Введение в теорию вероятностей

Геофизические измерения являются результатом случайных событий и случайных экспериментов. Целью геофизических экспериментов является определение как неизвестных параметров Земли, так и их доверительных интервалов, проверка гипотез относительно существования тех или иных значений параметров или состояний объектов.

1.1 Определение вероятности

Основным понятием теории вероятностей является понятие случайного события. Случайное событие появляется в результате измерения или эксперимента. Определение вероятностей случайных событий составляет предмет теории вероятностей.

1.1.1 Множество элементарных событий

С данным опытом связано множество Ω элементарных событий ω_i ($i = 1, \dots, n$)

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\},$$

так что в результате опыта появляется одно и только одно элементарное событие.

Пример 1: Подбрасывание монеты: “орел и решетка” $\Omega = \{H, T\}$.

Пример 2: n подбрасываний монеты:

$$\Omega = \{\omega : \omega = (a_1, a_2, \dots, a_n), a_i = \{H, T\}\}.$$

Общее число исходов $N(\Omega) = 2^n$.

Операции над множествами

Множество A принадлежит множеству Ω если каждый элемент множества A является элементом множества Ω , $A \subset \Omega$ (см. рис. 1.1). Множество не содержащее элементов называется пустым множеством и обозначается \emptyset . В теории вероятностей множество \emptyset называется невозможным событием, а множество Ω — достоверным событием.

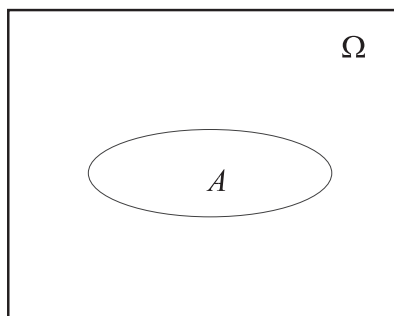


Рис. 1.1: Множество A является подмножеством множества Ω .

Равенство множеств. Множество A_1 равно множеству A_2 если, и только если, каждый элемент из A_1 является элементом A_2 и каждый элемент из A_2 является элементом из A_1 (см. рис. 1.2):

$$A_1 = A_2 \quad (\quad A_1 \subset A_2 \quad \text{и} \quad A_2 \subset A_1.)$$

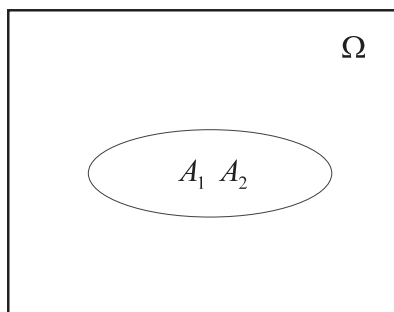


Рис. 1.2: Множество A_1 равно множеству A_2 .

Произведение множеств. Произведением или пересечением

$$A_1 \cdot A_2 = A_1 \cap A_2 = \{ \omega \in \Omega : \omega \in A_1 \text{ and } \omega \in A_2 \}$$

двух множеств A_1 и A_2 является множество, состоящее из общих для множества A_1 и множества A_2 элементов (см. рис. 1.3)

$$A_1 \cdot A_2 = A_2 \cdot A_1, \quad A_1 \cdot \emptyset, \quad A_1 \cdot \Omega = A_1.$$

Множества A_1 и A_2 называются несовместными (или непересекающимися), если они не имеют общих элементов

$$A_1 \cdot A_2 = \emptyset.$$

Множества A_1, A_2, \dots называются несовместными, если любые два множества не имеют общих элементов

$$A_i \cdot A_j = \emptyset \text{ for every } i, j, \quad i \neq j.$$

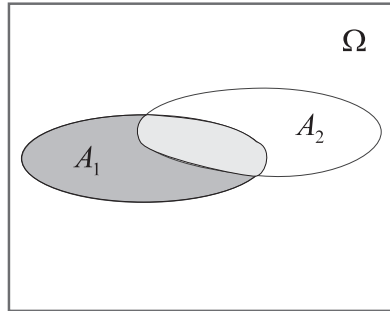


Рис. 1.3: Пересечение двух множеств.

Объединение. Объединением

$$A_1 \cup A_2 = \{\omega \in \Omega : \omega \in A_1 \text{ или } \omega \in A_2\}$$

двух множеств A_1 и A_2 называют множество, состоящее из элементов, принадлежащих множеству A_1 и множеству A_2 , включая общие для этих множеств элементы (см. рис. 1.4). Справедливы следующие равенства

$$A_1 \cup A_2 = A_2 \cup A_1, \quad A_1 \cup A_1 = A_1, \quad A_1 \cup \emptyset = A_1, \quad A_1 \cup \Omega = \Omega.$$

Если $A_1 \cap A_2 = \emptyset$, то, в случае операции объединения, вместо знака \cup используют знак $+$ и называют объединение суммой.

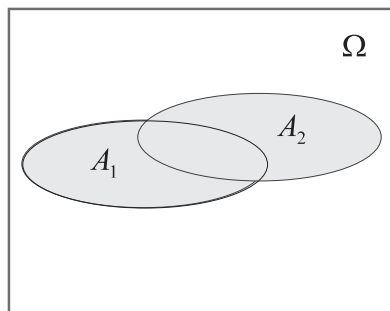


Рис. 1.4: Объединение двух множеств.

Дополнение. Дополнение \bar{A} множества A определяется как множество, состоящее из всех элементов множества Ω , которые не принадлежат множеству A (см. рис. 1.5). Справедливы следующие равенства

$$\begin{aligned}\bar{A} &= \Omega \setminus A = \Omega - A, & \bar{A} \cdot A &= \bar{A} \cap A = \emptyset, \\ \bar{\emptyset} &= \Omega, & \bar{\Omega} &= \emptyset, & A + \bar{A} &= \Omega.\end{aligned}$$

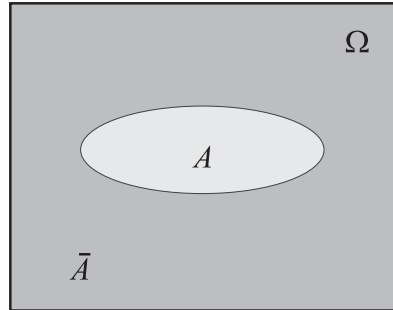


Рис. 1.5: Дополнением множества A является множество \bar{A} .

Разность. Разностью $A_1 - A_2$ множеств A_1 и A_2 называют множество состоящее из элементов множества A_1 , которые не являются элементами множества A_2 (см. рис. 1.6). Справедливы следующие равенства

$$A - \emptyset = A, \quad A - \bar{A} = \emptyset, \quad \Omega - A = \bar{A}.$$

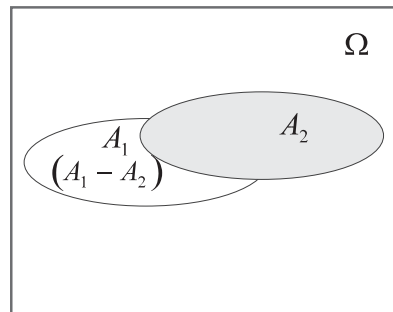


Рис. 1.6: Разность двух множеств.

Закон Моргана

Легко показать, что

$$\overline{A_1 \cup A_2} = \bar{A}_1 \cdot \bar{A}_2, \quad \overline{(A_1 \cdot A_2)} = \bar{A}_1 \cup \bar{A}_2.$$

В случае равенства множеств мы заменяем сумму произведением, произведение — суммой, и множества их дополнением. Равенство сохраняется (см. рис. 1.7). Множество Ω называют достоверным событием. Пустое множество \emptyset называют

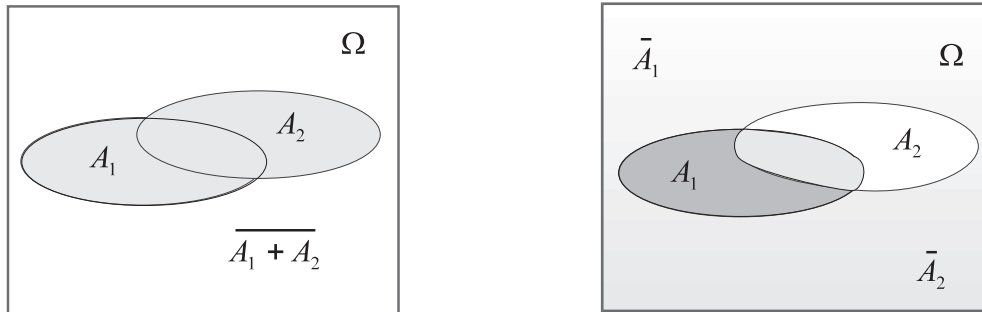


Рис. 1.7: Иллюстрация закона Моргана.

невозможным событием.

Алгебра событий

Множество \mathcal{A} , удовлетворяющее условиям:

1. $\Omega \in \mathcal{A}$.
2. $\emptyset \in \mathcal{A}$.
3. Если $A \in \mathcal{A}$, то $\bar{A} \in \mathcal{A}$.
4. Если $A_1 \in \mathcal{A}$ и $A_2 \in \mathcal{A}$, то $A_1 \cup A_2 \in \mathcal{A}$ и $A_1 \cap A_2 \in \mathcal{A}$; то \mathcal{A} называется алгеброй событий.

1.1.2 Аксиоматическое определение вероятности.

1. Пространство элементарных событий — Ω .
2. Множество событий — алгебра событий \mathcal{A} .
3. Вероятность исходов (Аксиомы вероятности).

Аксиома 1: Вещественное число $P(\omega_i) \geq 0$, задано для каждого события $\omega_i \in \Omega$ ($i = 1, 2, \dots, n$). $P(\omega_i)$ называется вероятностью события ω_i .

Аксиома 2: Вероятность достоверного события равна единице, $P(\Omega) = 1$.

Аксиома 3: Если $\omega_1, \omega_2, \dots$ есть счетная последовательность непересекающихся $\omega_i \cap \omega_j = \emptyset$ for $i \neq j$ событий, принадлежащая Ω , то

$$P(\omega_1 \cup \omega_2 \cup \dots) = P(\omega_1) + P(\omega_2) + \dots$$

Нормировка: $p(\omega_1) + p(\omega_2) + \dots + p(\omega_n) = 1$.

Вероятность события $A \in \mathcal{A}$ есть

$$P(A) = \sum_{\{\omega_i \in A\}} p(\omega_i).$$

Совокупность объектов (Ω, \mathcal{A}, P) называют полем вероятностей

1.1.3 Основные свойства вероятности

Из рассмотренных выше аксиом можно получить:

1. $P(\emptyset) = 0$.
2. $P(\Omega) = 1$.
3. $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$.
Если $(A_1 \cap A_2) = \emptyset$, то

$$P(A_1 + A_2) = P(A_1) + P(A_2).$$

Если A_1, A_2, \dots, A_n взаимно независимы, то

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

В общем случае:

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{i=1}^n \sum_{j=1}^n P(A_i A_j) + \\ &+ \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n P(A_i A_j A_k) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n). \end{aligned}$$

4. Если $A_1 \subset A_2$, то $P(A_1) \leq P(A_2)$, $P(A_2) = P(A_1) + P(A_2 \bar{A}_1) \geq P(A_1)$.
5. $P(\bar{A}) = 1 - P(A)$, $A + \bar{A} = \Omega$, $A\bar{A} = \emptyset$,
 $P(A + \bar{A}) = P(A) + P(\bar{A}) - P(A\bar{A}) = 1$.
6. Для произвольных событий A_1, A_2, \dots, A_n справедливо неравенство

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i).$$

1.1.4 Частотное определение вероятности

Эксперимент повторяется n раз. Если событие A реализуется n_A раз, тогда вероятность $P(A)$ события n_A определяется как предел относительной частоты n_A/n реализации события A :

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}.$$

Например, если монета подбрасывается n раз и “орел” выпадает n_h раз, то вероятность выпадения “орла” есть предел n_h/n .

1.1.5 Классическое определение

Вероятность $P(A)$ события A находится “априори” без проведения эксперимента путем подсчета общего числа n возможных исходов. Если событие A появляется n_A раз, то вероятность $P(A)$ дается формулой

$$P(A) = \frac{n_A}{n},$$

где n_A есть число благоприятных исходов для события A . Вероятность $P(A)$ равна отношению благоприятных исходов к общему числу исходов.

Пример: При однократном подбрасывании игральной кости возможными исходами являются 1, 2, 3, 4, 5, 6. Вероятность выпадения “четных” чисел 2, 4, 6 равна $P(\text{“четные”}) = 3/6$.

1.1.6 Геометрическое распределение

Важным случаем множества событий является вещественная ось или часть вещественной оси.

В общем случае, подобные множества являются геометрическими объектами различной мерности (1-D, 2-D и 3-D) (см. рис. 1.8).

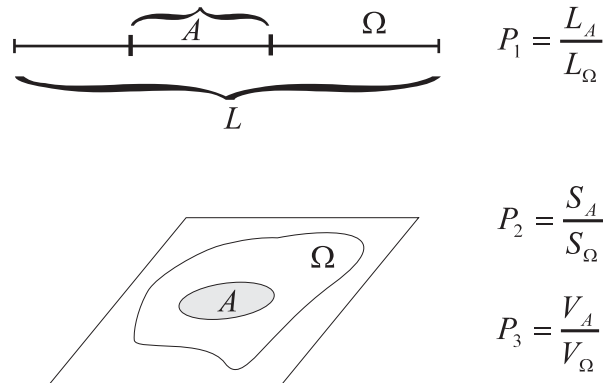


Рис. 1.8: Геометрическое определение вероятности.

1.1.7 Условная вероятность и независимость случайных событий

Иногда необходимо рассматривать вероятность некоторого события при условии, что произошло некоторое событие

Пусть вероятность $P(A_2)$ случайного события A_2 не равна нулю, тогда *условная вероятность* $P(A_1/A_2)$ события A_1 , при условии реализации события A_2 дается отношением

$$P(A_1/A_2) = \frac{P(A_1A_2)}{P(A_2)} \quad \text{with } P(A_2) > 0,$$

при этом

$$P(A_1A_2) = P(A_2)P(A_1/A_2).$$

Если событие A_1 не зависит от события A_2 , тогда

$$P(A_1/A_2) = P(A_1)$$

и

$$P(A_1A_2) = P(A_1)P(A_2).$$

Для независимы событий A_1, A_2, \dots, A_n справедливо

$$P(A_1A_2 \dots A_n) = P(A_1)P(A_2) \dots P(A_n).$$

1.1.8 Формула полной вероятности

Пусть имеется n взаимно независимых событий

$$B_1, B_2, \dots, B_n,$$

сума которых равна достоверному событию Ω

$$\begin{aligned} B_i B_j &= \emptyset, \quad i \neq j, \quad i, j = 1, 2, \dots, n, \\ B_1 + B_2 + \dots + B_n &= \Omega. \end{aligned}$$

Утверждение: Пусть A — произвольное событие. Можно показать, что

$$P(A) = P(A/B_1)P(B_1) + \dots + P(A/B_n)P(B_n).$$

Доказательство: Т.к. $A = A\Omega$, то из

$$A = A(B_1 + \dots + B_n) = AB_1 + \dots + AB_n \quad \text{and} \quad (AB_i)(AB_j) = \emptyset$$

можно получить следующее равенство

$$P(A) = P(AB_1) + \dots + P(AB_n).$$

Учитывая

$$P(AB_i) = P(A/B_i)P(B_i),$$

получим окончательно

$$P(A) = \sum_{i=1}^n P(A/B_i)P(B_i).$$

1.1.9 Теорема Байеса

Покажем, что

$$P(B_i/A) = \frac{P(A/B_i)P(B_i)}{\sum_{i=1}^n P(B_i)P(A/B_i)},$$

где события B_i удовлетворяют условиям $B_i B_j = \emptyset$ если $B_i \neq B_j$, $\sum_{i=1}^n B_i = \Omega$. Это формула приводит к определению *апостериорной вероятности* $P(B_i/A)$ события B_i через (*априорную*) вероятность $P(B_i)$ и условную вероятность $P(A/B_i)$.

Доказательство: Пусть мы имеем

$$P(B_i A) = P(A)P(B_i/A) = P(B_i)P(A/B_i).$$

Тогда

$$P(B_i/A) = \frac{P(B_i)P(A/B_i)}{P(A)}.$$

1.2 Функция распределения

1.2.1 Случайная переменная

Вещественная функция $\xi = \xi(\omega_i)$, определенная на поле вероятностей (Ω, \mathcal{A}, P) на множестве Ω элементарных событий ω_i называется *случайной переменной*, если для каждого $\xi \in \mathcal{R}$ событие, задаваемое неравенством $\xi(\omega_i) < x$, принадлежит случайному событию из \mathcal{A} .

Пример: Подбрасывание монеты: $\omega_1 = HH$, $\omega_2 = HT$, $\omega_3 = TH$, $\omega_4 = TT$, $\xi(\omega_1) = 2$, $\xi(\omega_2) = 2$, $\xi(\omega_3) = 1$, $\xi(\omega_4) = 0$.

$$P_\xi(x_i) = P\{\omega : \xi(\omega) = x_i\}, \quad x_i \in X.$$

Дискретная функция $\{P_\xi(x_1), \dots, P_\xi(x_n)\}$ называется *распределением вероятностей* случайной переменной ξ (см. рис. 1.9).

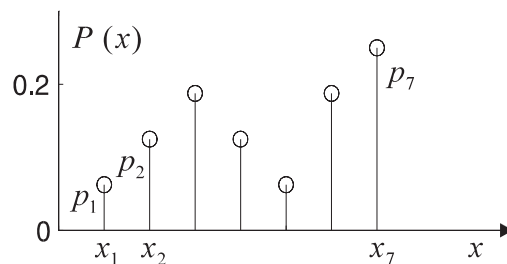


Рис. 1.9: Графическое представление распределения вероятностей.

1.2.2 Определение функции распределения

Для вещественного числа x , множество $\{\xi(\omega) \leq x\}$ состоит из всех событий ω , таких, что $\xi(\omega) \leq x$. При этом вероятность таких событий $P\{\xi(\omega) \leq x\}$ зависит от x , т. е. является функцией от x . Обозначая такую функцию как

$$F_\xi(x) = P\{\omega : \xi(\omega) \leq x\}.$$

Функция $F_\xi(x)$, определенная на всей вещественной оси x от $-\infty$ до ∞ , называется *функцией распределения* случайной переменной ξ (см. рис. 1.10).

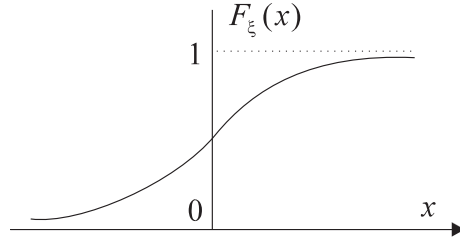


Рис. 1.10: Пример функции распределения (непрерывная случайная величина).

Функция распределения дискретной случайной величины определяется следующим образом (см. рис. 1.11)

$$F_\xi(x) = \sum_{\{i: x_i \leq x\}} P_\xi(x_i),$$

где

$$P_\xi(x_i) = F_\xi(x_i) - F_\xi(x_i^-), \quad F_\xi(x_i^-) = \lim_{y \rightarrow x_i^-} F_\xi(y)$$

и для $x_1 < x_2 < \dots < x_{i-1} < x_i$

$$F_\xi(x_0)|_{x_0 < x_1} = 0, \quad P_\xi(x_i) = F_\xi(x_i) - F_\xi(x_{i-1}).$$

1.2.3 Свойства функции распределения

Функция распределения $F(x)$ обладает следующими свойствами:

1. $F_\xi(-\infty) = 0, F_\xi(\infty) = 1$.
2. Функция распределения — невозрастающая функция:

$$F(x_1) \leq F(x_2) \quad \text{for } x_1 < x_2.$$

3. Если $\alpha < \beta$, то

$$P_\xi(\alpha < \xi \leq \beta) = F_\xi(\beta) - F_\xi(\alpha).$$

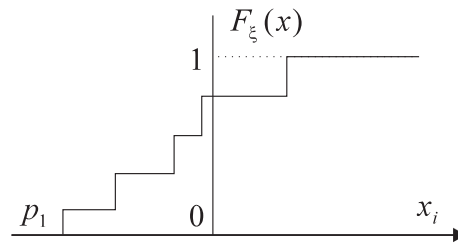


Рис. 1.11: Пример функции распределения (непрерывная случайная величина).

1.2.4 Плотность вероятностей

Производная от

$$f_{\xi}(x) = \frac{dF_{\xi}(x)}{dx}$$

функции распределения $F_{\xi}(x)$ называется *плотностью вероятностей* случайной переменной ξ .

Свойства плотности распределения вероятностей.

1. Из монотонности $F_{\xi}(x)$ следует неотрицательность $f_{\xi}(x)$:

$$f_{\xi}(x) \geq 0.$$

2. Функция распределения вероятностей есть интеграл от плотности распределения вероятностей

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(y) dy,$$

следовательно

$$F_{\xi}(x_2) - F_{\xi}(x_1) = \int_{x_1}^{x_2} f_{\xi}(x) dx.$$

- 3.

$$\int_{-\infty}^{\infty} f_{\xi}(x) dx = F_{\xi}(\infty) - F_{\xi}(-\infty) = 1.$$

4. Вероятность события $x_1 < \xi \leq x_2$ есть интеграл от плотности вероятностей (см. рис. 1.12)

$$P(x_1 < \xi \leq x_2) = \int_{x_1}^{x_2} f_{\xi}(x) dx.$$

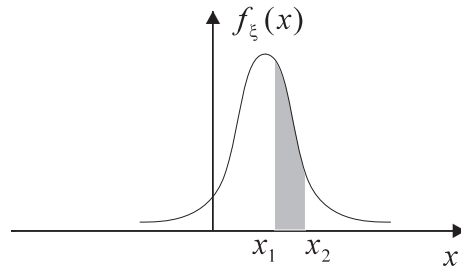


Рис. 1.12: Вероятность есть интеграл от плотности вероятностей.

Если Δx достаточно малый интервал, то

$$P(x < \xi \leq x + \Delta x) \simeq f(x)\Delta x.$$

Плотность вероятности может быть определена как предел

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < \xi \leq x + \Delta x)}{\Delta x}.$$

Отметим, что $P(\xi = x) = 0$.

1.2.5 Функция распределения и плотность вероятностей для функции, зависящей от случайной переменной

Определим плотность вероятности $f_\eta(y)$ случайной переменной

$$\eta = \varphi(\xi)$$

в терминах плотности вероятности $f_\xi(x)$ of ξ . Функция $\varphi(\xi)$ непрерывно дифференцируема и монотонна относительно своего аргумента ξ .

Пусть ψ — обратная функция по отношению к функции φ , тогда плотность вероятности можно записать

$$f_\eta(y) = f_\xi(\psi(y))|\psi'(y)|.$$

Для $\eta < y$ $\{\Delta_1(y), \Delta_2(y), \dots, \Delta_n(y)\}$ функция распределения может быть представлена в виде (см. рис. 1.13)

$$\begin{aligned} F_\eta(y) &= P\{\omega : \eta(\omega) \leq y\} = \\ &= P\{\omega : (\xi(\omega) \in \Delta_1(y) + \xi(\omega) \in \Delta_2(y) + \dots)\} = \\ &= \sum_i P\{\omega : \xi(\omega) \in \Delta_i(y)\}. \end{aligned}$$

Пример: Если случайная переменная ξ имеет нормальное распределение

$$\xi \in N(m_\xi, \sigma_\xi^2), \quad f_\xi(x) = \frac{\exp\left(-\frac{(x-m_\xi)^2}{2\sigma_\xi^2}\right)}{\sqrt{2\pi}\sigma_\xi},$$

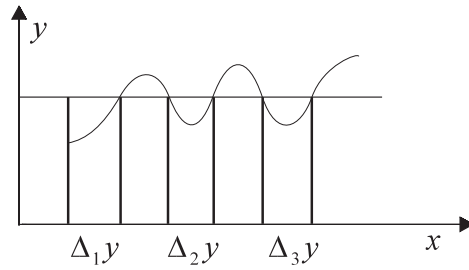


Рис. 1.13: Иллюстрация замены переменной.

тогда линейная функция $\eta = a_1 + a_2\xi$ имеет плотность вероятности:

$$f_\eta(y) = \frac{1}{|a_2|\sigma_\xi\sqrt{2\pi}} \exp\left(-\frac{[y - (a_2m_\xi + a_1)]^2}{2|a_2|^2\sigma_\xi^2}\right),$$

$$m_\eta = a_2m_\xi + a_1, \quad \sigma_\eta = |a_2|\sigma_\xi \quad (f_\eta(y) = N(m_\eta, \sigma_\eta^2)),$$

т.к.

$$f_\eta(y) = f_\xi(\psi(y))|\psi'(y)|, \quad \psi(y) = (y - a_1)/a_2, \quad |\psi'(y)| = 1/|a_2|.$$

1.2.6 Случайные векторы

Однозначная функция $\vec{\xi}(\omega_i) = (\xi_1(\omega_i), \dots, \xi_m(\omega_i))$ со значениями в \mathcal{R}^n , определенная на множестве $\text{set } \Omega$ (на поле вероятностей (Ω, \mathcal{A}, P)) элементарных событий ω_i называется m -мерным *случайным вектором*. Функция распределения случайного вектора определяется следующим образом

$$F_\xi(x_1, \dots, x_m) = P\{\omega : \xi_1(\omega) \leq x_1, \dots, \xi_m(\omega) \leq x_m\}.$$

Пусть даны две случайные переменные (ξ, η) . Совместная функция распределения случайных переменных ξ и η определяется как

$$F_{\xi\eta}(x, y) = P\{\omega : \xi(\omega) \leq x, \eta(\omega) \leq y\}.$$

Свойства функции распределения:

1. $x_2 > x_1, F_{\xi\eta}(x_2, y) \geq F_{\xi\eta}(x_1, y), \quad y_2 > y_1, F_{\xi\eta}(x, y_2) \geq F_{\xi\eta}(x, y_1)$.
2. $F_{\xi\eta}(x, -\infty) = F_{\xi\eta}(-\infty, y) = F_{\xi\eta}(-\infty, -\infty) = 0$.
3. $F_{\xi\eta}(x, +\infty) = F_\xi(x), F_{\xi\eta}(+\infty, y) = F_\eta(y), F_{\xi\eta}(+\infty, +\infty) = 1$.

Функция

$$f_{\xi\eta}(x, y) = \frac{\partial^2 F_{\xi\eta}(x, y)}{\partial x \partial y}$$

называется *совместной плотностью вероятности*. Вероятность события $(\xi(\omega), \eta(\omega) \in D)$ записывается в виде

$$P\{\omega : (\xi(\omega), \eta(\omega)) \in D\} = \iint_D f_{\xi\eta}(x, y) dx dy$$

и функция распределения может быть представлена как

$$F_{\xi\eta}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{\xi\eta}(p, q) dp dq.$$

Свойства плотности вероятности:

1. $f_{\xi\eta}(x, y) \geq 0$.
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\xi\eta}(x, y) dx dy = 1$.

1.2.7 Маргинальные и условные распределения

Маргинальная функция распределения —

$$F_{\xi}(x) = F_{\xi\eta}(x, \infty), \quad F_{\eta}(y) = F_{\xi\eta}(\infty, y),$$

$$F_{\xi}(x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(p, q) dp dq, \quad F_{\eta}(y) = \int_{-\infty}^{\infty} \int_{-\infty}^y f(p, q) dp dq$$

и *маргинальные плотности вероятности*:

$$f_{\xi}(x) = F'_{\xi}(x) = \int_{-\infty}^{\infty} f(x, y) dy, \quad f_{\eta}(y) = F'_{\eta}(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Условные плотности вероятности (см. рис. 1.14):

$$f_{\xi/\eta}(x/y) = \frac{f_{\xi\eta}(x, y)}{f_{\eta}(y)} = \frac{f_{\xi\eta}(x, y)}{\int_{-\infty}^{\infty} f_{\xi\eta}(x, y) dx},$$

$$f_{\eta/\xi}(y/x) = \frac{f_{\xi\eta}(x, y)}{f_{\xi}(x)} = \frac{f_{\xi\eta}(x, y)}{\int_{-\infty}^{\infty} f_{\xi\eta}(x, y) dy}$$

$$f_{\xi\eta}(x, y) = f_{\eta}(y) f_{\xi/\eta}(x/y) = f_{\xi}(x) f_{\eta/\xi}(y/x).$$

Две случайные величины ξ и η являются независимыми, если справедливы следующие соотношения

$$f_{\xi/\eta}(x/y) = f_{\xi}(x), \quad f_{\eta/\xi} = f_{\eta}(y), \quad f_{\xi\eta}(x, y) = f_{\xi}(x) f_{\eta}(y).$$

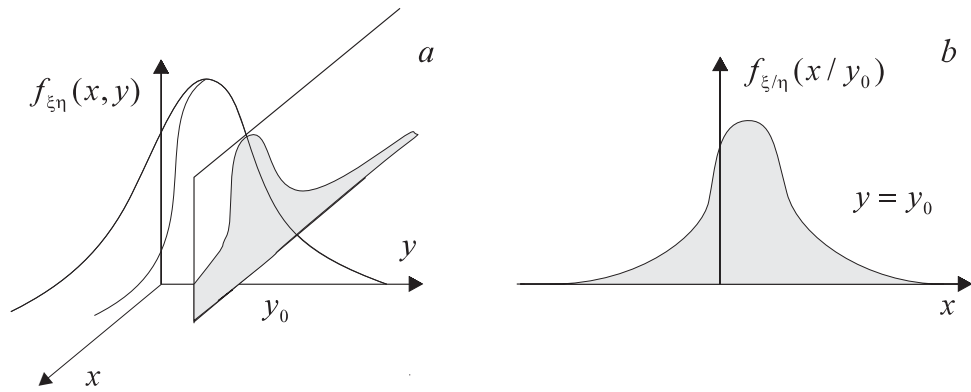


Рис. 1.14: Графическое представление совместной и условной плотностей вероятности. (а) – совместная плотность вероятности; (б) – условная плотность вероятности.

Аналог формулы Байеса:

$$f_{\xi/\eta}(x, y) = \frac{f_{\xi}(x) f_{\eta/\xi}(y/x)}{\int_{-\infty}^{\infty} f_{\xi\eta}(x, y) dx}.$$

1.2.8 Совместные распределения и плотности вероятностей

Две случайные переменные (ξ, η) описываются плотностью вероятностей $f_{\xi\eta}(x, y)$ и случайная переменная $\zeta = \varphi(\xi, \eta)$ имеет функцию распределения $F_{\zeta}(z)$ (см. рис. 1.15):

$$F_{\zeta}(z) = P\{\omega : \zeta(\omega) < z\} = P\{\omega : \varphi(\xi(\omega), \eta(\omega)) < z\},$$

$$F_{\zeta}(z) = P\{\omega : (\xi, \eta) \in \mathcal{D}\} = \iint_{\mathcal{D}(z)} f_{\xi\eta}(x, y) dx dy.$$

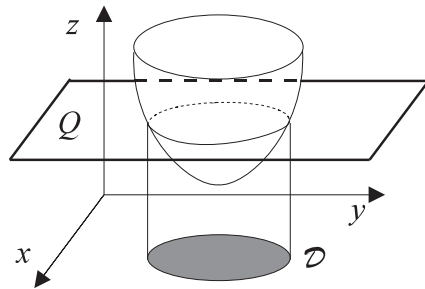


Рис. 1.15: Иллюстрация вычисления совместной функции распределения.

Пример 1: Рассмотрим важный пример функции двух случайных переменных. Пусть $\zeta = \xi + \eta$. В случае определения $F_{\zeta}(z)$, область $\mathcal{D}(z)$ на плоскости

(x, y) , задается как

$$x + y < z,$$

(см. рис. 1.16), т.е. является полуплоскостью, расположенной слева от линии $x + y = z$. Дифференцируя функцию распределения

$$\begin{aligned} F_{\zeta}(z) &= \iint_{\mathcal{D}(z)} f_{\xi\eta}(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_{\xi\eta}(x, y) dx dy = \\ &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{z-x} f_{\xi\eta}(x, y) dy \right\} dx \end{aligned}$$

по z , получим

$$f_{\zeta}(z) = \int_{-\infty}^{\infty} f_{\xi\eta}(x, z-x) dx \quad \text{or} \quad f_{\zeta}(z) = \int_{-\infty}^{\infty} f_{\xi\eta}(z-y, y) dy.$$

Если ξ и η — независимые случайные переменные, тогда плотность веро-

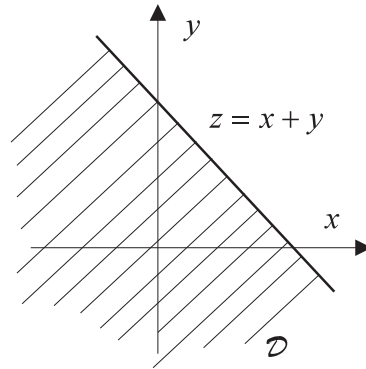


Рис. 1.16: Область интегрирования на плоскости (x, y) .

ятности их суммы $\zeta = \xi + \eta$ равна свертке их плотностей вероятностей

$$f_{\zeta}(z) = \int_{-\infty}^{\infty} f_{\xi}(x) f_{\eta}(z-x) dx, \quad f_{\zeta}(z) = \int_{-\infty}^{\infty} f_{\xi}(z-y) f_{\eta}(y) dy,$$

or

$$f_{\zeta}(z) = f_{\xi} * f_{\eta}.$$

Пример 2: Предположим, что случайные переменные ξ и η имеют нормальные распределения

$$\begin{aligned} \xi &\in N(m_{\xi}, \sigma_{\xi}^2), & f_{\xi}(x) &= \frac{1}{\sigma_{\xi} \sqrt{2\pi}} \exp\left(-\frac{(x - m_{\xi})^2}{2\sigma_{\xi}^2}\right), \\ \eta &\in N(m_{\eta}, \sigma_{\eta}^2), & f_{\eta}(y) &= \frac{1}{\sigma_{\eta} \sqrt{2\pi}} \exp\left(-\frac{(y - m_{\eta})^2}{2\sigma_{\eta}^2}\right). \end{aligned}$$

Пусть $\zeta = \xi + \eta$. Найдем функцию $f_\zeta(z)$.

Плотность вероятности $f_\zeta(z)$ является сверткой

$$f_\zeta(z) = \int_{-\infty}^{\infty} f_\xi(x) f_\eta(z-x) dx,$$

$$f_\zeta(z) = \frac{1}{2\pi\sigma_\xi\sigma_\eta} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-m_\xi)^2}{2\sigma_\xi^2} - \frac{(z-x-m_\eta)^2}{2\sigma_\eta^2}\right) dx,$$

$$f_\zeta(z) = \frac{1}{2\pi\sigma_\xi\sigma_\eta} \int_{-\infty}^{\infty} \exp(-a_2x^2 + 2a_1x - a_0) dx,$$

$$a_2 = \frac{1}{2} \frac{\sigma_\xi^2 + \sigma_\eta^2}{\sigma_\xi\sigma_\eta}, \quad a_1 = \frac{m_\xi}{2\sigma_\xi^2} + \frac{z-m_\eta}{2\sigma_\eta^2}, \quad a_0 = \frac{m_\xi^2}{2\sigma_\xi^2} + \frac{(z-m_\eta)^2}{2\sigma_\eta^2},$$

$$\int_{-\infty}^{\infty} \exp(-a_2x^2 + 2a_1x - a_0) dx = \sqrt{\frac{\pi}{a_2}} \exp\left(-\frac{a_2a_0 - a_1^2}{a_2}\right)$$

и получим

$$f_\zeta(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\xi^2 + \sigma_\eta^2}} \exp\left(-\frac{[z - (m_\xi + m_\eta)]^2}{2(\sigma_\xi^2 + \sigma_\eta^2)}\right),$$

т.е.

$$\zeta \in N(m_\zeta, \sigma_\zeta^2), \quad m_\zeta = m_\xi + m_\eta, \quad \sigma_\zeta^2 = \sigma_\xi^2 + \sigma_\eta^2.$$

1.3 Математическое ожидание и моменты случайных величин

Пусть задано поле вероятностей (Ω, \mathcal{A}, P) и случайная величина $\xi(\omega)$ принимает значения $X = \{x_1, \dots, x_n\}$. Связь между случайным событием $A_i = \{\omega : \xi(\omega) = x_i\}$ ($i = 1, \dots, n$) и значением случайной величины $\xi(\omega)$ можно записать следующим образом

$$\xi(\omega) = \sum_{i=1}^n x_i I(A_i); \quad I(A_i) = \begin{cases} 1 & \omega \in A_i, \\ 0 & \omega \notin A_i. \end{cases}$$

1.3.1 Математическое ожидание

Математическое ожидание есть результат усреднения всех возможных значений случайной переменной с учетом их вероятностей. Если ξ дискретная случайная величина, усреднение x_i с вероятностью $P(A_i)$ осуществляется по формуле

as

$$M\xi = \sum_{i=1}^n x_i P(A_i), \quad A_i = \{\omega : \xi(\omega) = x_i\}, \quad P_\xi(x_i) = P(A_i),$$

$$M\xi = \sum_{i=1}^n x_i P_\xi(x_i); \quad \xi \in R^1.$$

Математическое ожидание непрерывной случайной величины ξ дается интегралом вида

$$M\xi = \int_{-\infty}^{\infty} x f_\xi(x) dx,$$

где $f_\xi(x)$ — плотность вероятности ξ (см. рис. 1.17).

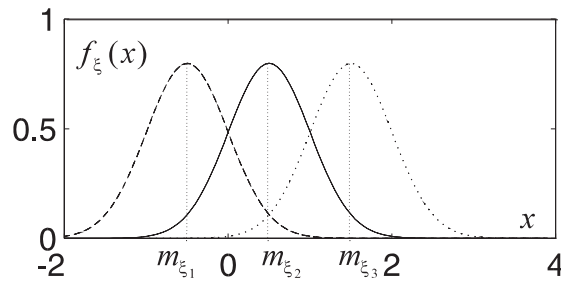


Рис. 1.17: Пример графического представления нормальной плотности вероятности с различными величинами математического ожидания. Среднеквадратическое отклонение: $\sigma_{\xi_1} = \sigma_{\xi_2} = \sigma_{\xi_3} = 0.5$. Математическое ожидание: $m_{\xi_1} = -0.5$, $m_{\xi_2} = 0.5$, $m_{\xi_3} = 1.5$.

Свойства математического ожидания

- $\xi \geq 0, \quad M\xi \geq 0.$
- $M(a\xi + b\eta) = aM\xi + bM\eta.$
- $\xi \geq \eta, \quad M\xi \geq M\eta.$
- $|M\xi| \leq M|\xi|.$
- Если ξ and η — независимые случайные величины, тогда $M(\xi \cdot \eta) = M\xi \cdot M\eta.$
- $(M|\xi\eta|)^2 \leq M\xi^2 \cdot M\eta^2$ — неравенство Коши-Буняковского.
- $M\varphi(\xi) = \sum_i \varphi(x_i) P_\xi(x_i), \quad M\varphi(\xi) = \int_{-\infty}^{\infty} \varphi(x) f_\xi(x) dx.$

1.3.2 Дисперсия

Аналогом математического ожидания $M\xi$ случайной величины является центр масс (центр тяжести) “некоторого тела” с функцией плотности $f(x)$. Другая важная характеристика случайной величины — *дисперсия* σ^2 определяется формулой вида

$$\xi(\omega) : D\xi = M(\xi - M\xi)^2, \quad \sigma_\xi = \sqrt{D\xi}.$$

Аналогом дисперсии является момент инерции механической системы, обладающей “вероятностной массой”. Квадратный корень из дисперсии обозначают σ_ξ и называют *среднеквадратическим* или *стандартным* отклонением (пример некоррелированного временного ряда с нормальным распределением и различными среднеквадратическими отклонениями дан на рис. 1.18). Если ξ дискретная случайная величина, то

$$\sigma_\xi^2 = \sum_i (x_i - M\xi)^2 P_\xi(x_i).$$

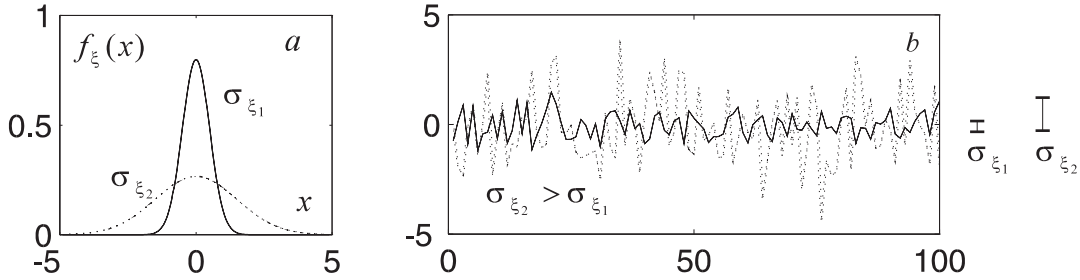


Рис. 1.18: Пример некоррелированных временных рядов с нормальным распределением при двух различных значениях среднеквадратического отклонения (b) и соответствующие им плотности вероятностей (a). Среднеквадратические отклонения: $\sigma_{\xi_1} = 0.5$ $\sigma_{\xi_2} = 1.5$. Математическое ожидание: $m_{\xi_1} = m_{\xi_2} = 0$.

1.3.3 Свойства дисперсии и коэффициентов корреляции

- $D\xi = M\xi^2 - (M\xi)^2$,
 $D\xi = M(\xi - M\xi)^2 = M\xi^2 - 2(M\xi)^2 + (M\xi)^2 = M\xi^2 - (M\xi)^2$.
- $D(\xi) \geq 0$.
- $D(a_1 + a_2\xi) = a_2^2 D\xi$,
 $D(a_1 + a_2\xi) = M(a_1 + a_2\xi - a_1 - a_2M\xi)^2 = a_2^2 M(\xi - M\xi)^2 = a_2^2 D\xi$.
- $D(\xi + \eta) = D\xi + D\eta + 2R_{\xi\eta}$,
 $D(\xi + \eta) = M((\xi - M\xi) + (\eta - M\eta))^2 =$
 $= M(\xi - M\xi)^2 + M(\eta - M\eta)^2 + 2M(\xi - M\xi)(\eta - M\eta) = D\xi + D\eta + 2R_{\xi\eta}$,
 где $R_{\xi\eta} = \text{cov}(\xi, \eta) = M(\xi - M\xi)(\eta - M\eta)$ — коэффициент ковариации and

$r(\xi, \eta) = r_{\xi\eta} = \text{cov}(\xi, \eta) / (\sqrt{D\xi}\sqrt{D\eta}) = R_{\xi\eta} / \sigma_\xi \sigma_\eta$ — коэффициент корреляции ($-1 \leq r_{\xi\eta} \leq 1$) (см. рис. 1.19).

Пусть с.в. η является линейной функцией с.в. ξ

$$\eta = a_1 \xi + a_2.$$

По определению коэффициента корреляции, можно записать

$$\begin{aligned} r_{\xi\eta} &= M(\xi - M\xi)(a_1\xi + a_2 - a_1M\xi - a_2) / \sigma_\xi |a_1| \sigma_\xi = \\ &= \frac{a_1}{|a_1|} \frac{D\xi}{\sigma_\xi^2} = \frac{a_1}{|a_1|}. \end{aligned}$$

Если $a_1 > 0$, то $r_{\xi\eta} = 1$. Если $a_1 < 0$ то $r_{\xi\eta} = -1$.

Если с.в. ξ и η — независимы, то

$\text{cov}(\xi, \eta) = M(\xi - M\xi) \cdot M(\eta - M\eta) = (M\xi - M\xi)(M\eta - M\eta) = 0$,
то в этом случае $D(\xi + \eta) = D\xi + D\eta$.

- $M(\xi \cdot \eta) = M\xi \cdot M\eta + R_{\xi\eta}$,
 $R_{\xi\eta} = M(\xi - M\xi)(\eta - M\eta) = M(\xi\eta) - M\xi M\eta - M\eta M\xi + M\xi M\eta =$
 $= M(\xi\eta) - M\xi M\eta$. Если с.в. ξ и η независимы, то $M(\xi\eta) = M\xi M\eta$.

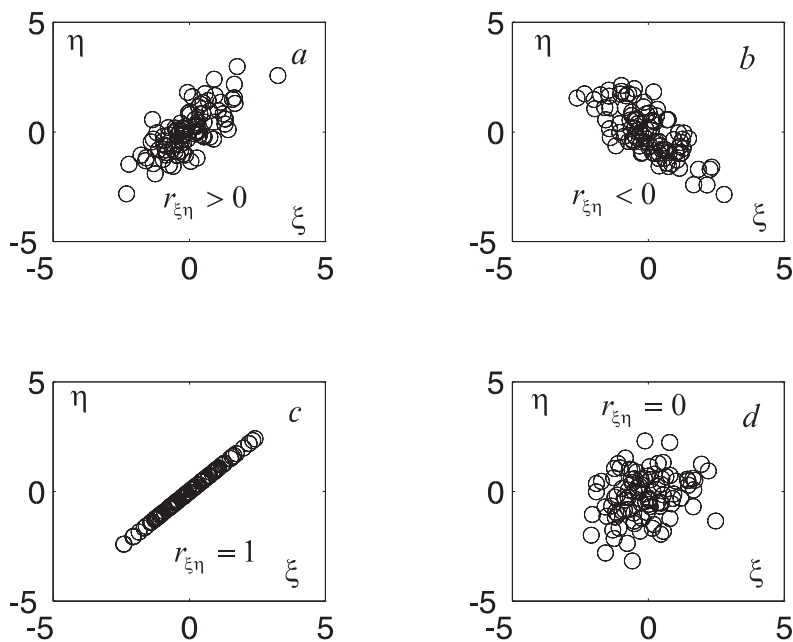


Рис. 1.19: Графическое представление выборки из двумерного нормального распределения с.в. ξ и η . $M\xi = M\eta = 0$; $r_{\xi\eta} = 0.75$ — (a); $r_{\xi\eta} = -0.75$ — (b); $r_{\xi\eta} = 1$ — (c); $r_{\xi\eta} = 0$ — (d). Размер выборки — 100.

$$\bullet D\left(\sum_{i=1}^n a_i \xi_i + a_0\right) = \sum_{i=1}^n a_i^2 D\xi_i + 2 \sum_{i<j} a_i a_j R_{\xi_i, \xi_j},$$

$$D\left(\sum_{i=1}^n a_i \xi_i + a_0\right) = M\left(\sum_{i=1}^n a_i \xi_i + a_0 - \sum_{i=1}^n a_i M(\xi_i) + a_0\right)^2 =$$

$$= M\left(\sum_{i=1}^n a_i (\xi_i - M\xi_i)\right)^2 = \sum_{i=1}^n a_i^2 D\xi_i + 2 \sum_{i<j} a_i a_j R_{\xi_i, \xi_j}.$$

Если с.в. ξ_i и ξ_j некоррелированы, то $D\left(\sum_{i=1}^n a_i \xi_i + a_0\right) = \sum_{i=1}^n a_i^2 D\xi_i$.

$$\bullet \text{ Пусть } \xi \text{ и } \eta \text{ — независимы с.в. } D(\xi\eta) = D\xi D\eta + (M\xi)^2 D\eta + (M\eta)^2 D\xi,$$

$$D(\xi\eta) = M(\xi\eta - M\xi M\eta)^2 = M(\xi\eta - M\xi M\eta)^2 =$$

$$= M(\xi^2 \eta^2) - 2M(\xi\eta) M\xi M\eta + (M\xi)^2 (M\eta)^2 = M\xi^2 M\eta^2 - (M\xi)^2 (M\eta)^2 =$$

$$= (D\xi + (M\xi)^2)(D\eta + (M\eta)^2) - (M\xi)^2 (M\eta)^2 = D\xi D\eta + (M\xi)^2 D\eta + (M\eta)^2 D\xi.$$

1.3.4 Квантили

Величина x_p , определяемая выражением

$$P(\xi < x_p) = F_\xi(x_p) = p, \quad 0 < p < 1$$

называется *нижней p -процентной точкой* или *p -ой квантилью* распределения (см. рис. 1.20). Соответственной, *верхняя p -процентная точка* определяется как

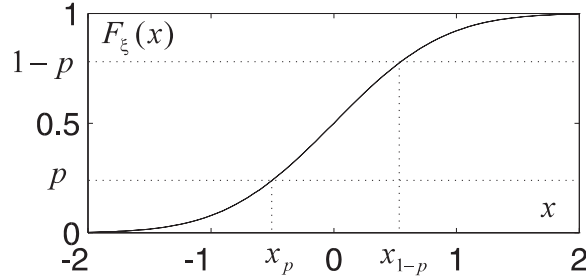


Рис. 1.20: Функция распределения p -процентная точка. Нормальное распределение:
 $m_\xi = 0, \sigma_\xi = 1$.

$$F_\xi(x_{1-p}) = 1 - p = 1 - F_x(x_p).$$

Величина $x_{1/2}$ называется *медианой* (см. рис. 1.21). Величины $x_{1/4}, x_{1/2}, x_{3/4}$ называются *квартелями* и $x_{0.1}, x_{0.2}, \dots, x_{0.9}$ называются *децилями* (см. рис. 1.22).

1.3.5 Характеристики плотности вероятности

Точка x_{max} , в которой плотность вероятности достигает своего максимального значения называется *модой* (см. рис. 1.23).

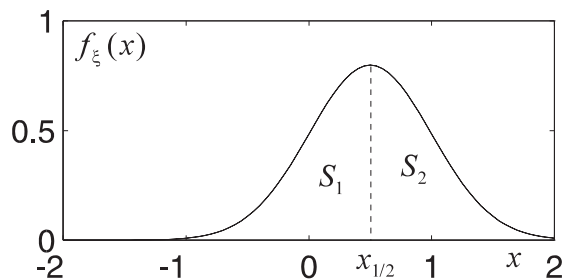


Рис. 1.21: Плотность вероятности и медиана. Нормальное распределение: $m_\xi = 0.5$, $\sigma_\xi = 0.5$.

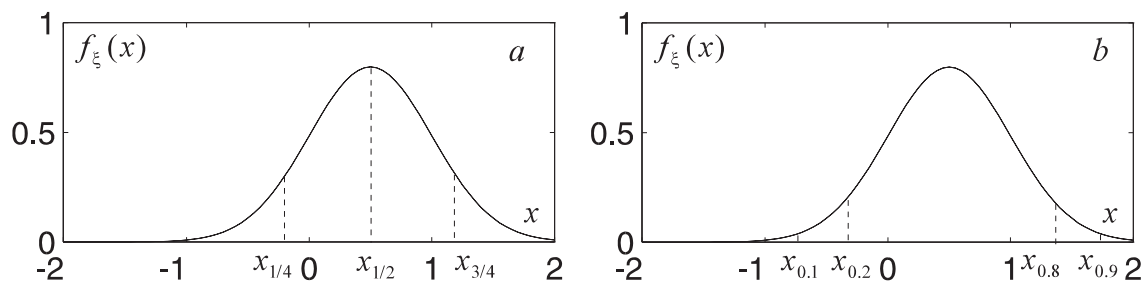


Рис. 1.22: Плотность вероятности и ее числовые характеристики: квантили (а) и децили (b). Нормальное распределение: $m_\xi = 0.5$, $\sigma_\xi = 0.5$.

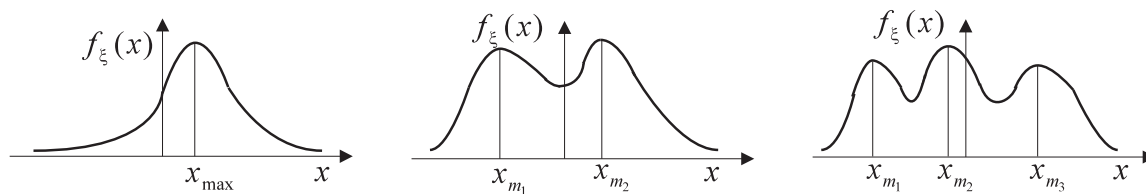


Рис. 1.23: Разномодальные распределения.

Отношение $\sigma_\xi/M\xi$ называется *коэффициентом вариации*. Величина $M|\xi - M\xi|$ называется *средним абсолютным отклонением*. Разность квантилей $x_{3/4} - x_{1/4}$ называется *интерквартильной шириной* (см. рис. 1.24).

Модуль разности $|x_{max} - x_{min}|$ называется *размахом* $P_\xi(x_i)$ (см. рис. 1.25).

Моменты одномерного распределения вероятностей

Математическое ожидание с.в. ξ определяет только “центр масс” плотности вероятности. Математическое ожидание является *начальным моментом* первого порядка случайной величины ξ . Начальные моменты других порядков определя-

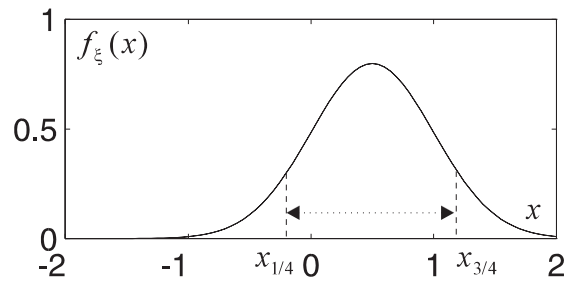


Рис. 1.24: Плотность вероятности и интерквартильная широта. Нормальное распределение: $m_\xi = 0.5$, $\sigma_\xi = 0.5$.

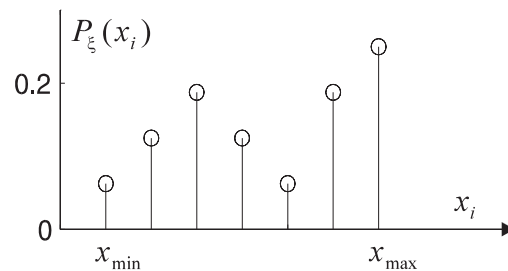


Рис. 1.25: Распределение вероятностей и его размах.

ются следующим образом

$$\alpha_s = M\xi^s = \int_{-\infty}^{\infty} x^s f(x) dx.$$

Ясно, что $\alpha_0 = 1$, $\alpha_1 = M\xi = m_\xi$. Величины

$$\mu_s = M(\xi - M\xi)^s = \int_{-\infty}^{\infty} (x - M\xi)^s f(x) dx$$

называются *центральными моментами*. Выполняются следующие равенства

$$\mu_0 = 1, \quad \mu_1 = 0, \quad \mu_2 = D\xi = \sigma^2 = \alpha_2 - m_\xi^2, \quad \mu_3 = \alpha_3 - 3m_\xi\alpha_2 + 2m_\xi^2.$$

Величина

$$\gamma_1 = \mu_3 / \sigma^3$$

называется *коэффициентом асимметрии* (см. рис. 1.26). Если плотность вероятности симметрична, то коэффициента γ_1 равен нулю. Величина

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$$

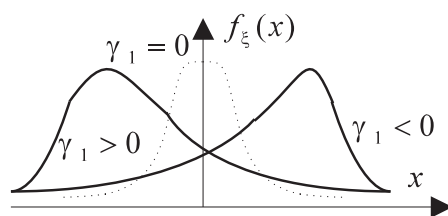


Рис. 1.26: Плотность вероятности и коэффициент асимметрии.

называется *эксцессом* (см. рис. 1.27). Эксцесс есть мера отклонения плотности вероятности $f_\xi(x)$ от нормального распределения. Если плотность вероятности принадлежит нормальному распределению, то эксцесс равен нулю.

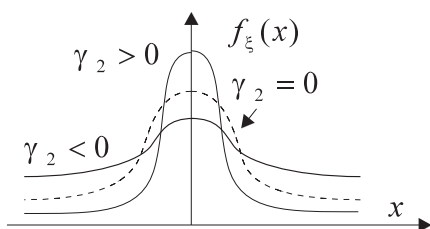


Рис. 1.27: Плотность вероятности и эксцесс.

1.4 Характеристические и производящие функции

1.4.1 Характеристические функции

Характеристическая функция $g(t)$ случайной величины ξ есть преобразование Фурье (с обратным знаком в показателе экспоненты) плотности вероятности $f_\xi(x)$.

Определение: *Характеристическая функция* с.в. ξ определяется как

$$g(t) = M(\exp(it\xi)).$$

Если ξ — непрерывная с.в., то

$$g(t) = \int_{-\infty}^{\infty} \exp(itx) f_\xi(x) dx$$

и, если ξ — дискретная с.в., то

$$g(t) = \sum_{k=1}^n \exp(itx_k) P_\xi(x_k).$$

Плотность вероятности $f_\xi(x)$ можно представить через $g(t)$ как интеграл

$$f_\xi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx)g(t)dt,$$

(обратное преобразование Фурье, формула обращения).

Свойства характеристических функций

1. Если ξ and η — с.в., такие, что $\eta = a\xi$, то $g_\eta(t) = g_\xi(at)$.

Доказательство: $g_\eta(t) = M \exp(it\eta) = M \exp(ita\xi) = M \exp(i(at)\xi) = g_\xi(at)$.

2. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые с.в. с характеристическими функциями $g_{\xi_1}(t), g_{\xi_2}(t), \dots, g_{\xi_n}(t)$ and $\eta = \sum_{k=1}^n \xi_k$, то $g_\eta(t) = \prod_{k=1}^n g_{\xi_k}(t)$.

Доказательство: $g_\eta(t) = M(\exp(it\eta)) = M(\exp(it \sum_{k=1}^n \xi_k)) =$
 $= M \left[\prod_{k=1}^n \exp(it\xi_k) \right] = \prod_{k=1}^n g_{\xi_k}(t)$.

3. Если ξ и η независимые случайные величины с плотностями вероятностей $f_\xi(x)$ и $f_\eta(y)$ соответственно, то плотность вероятностей и характеристическая функция случайной величины $\zeta = \xi + \eta$ равны:

$$f_\zeta(z) = f_\xi(x) * f_\eta(y)$$

и (используя преобразование Фурье)

$$g_\zeta(t) = g_\xi(t) \cdot g_\eta(t),$$

где

$$f_\zeta(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g_\zeta(t) \exp(-itz)dt.$$

1.4.2 Производящие функции моментов

Моменты случайных величин и их связь с плотностью вероятностей часто определяют при помощи *производящих функций моментов*. Пусть ξ — случайная величина с плотностью вероятностей $f_\xi(x)$. Тогда функцию $\alpha_\xi(t)$:

$$\alpha_\xi(t) = M(\exp(\xi t)) = \int_{-\infty}^{\infty} \exp(xt)f_\xi(x)dx$$

называют *производящей функцией моментов* случайной величины ξ .

Моменты случайной величины ξ могут быть получены из производящей функции моментов следующим образом. Раскладывая экспоненту в ряд Тейлора, получаем:

$$\alpha_\xi(t) = M\left[1 + \xi t + \frac{(\xi t)^2}{2!} + \frac{(\xi t)^3}{3!} + \dots\right] = 1 + \alpha_1 t + \alpha_2 \frac{t^2}{2!} + \alpha_3 \frac{t^3}{3!} + \dots$$

Производные функции $\alpha_s(t)$ связаны с моментами α_s . При $t = 0$ получаем

$$\frac{\partial^s \alpha_\xi(t)}{\partial t^s} = \int_{-\infty}^{\infty} x^s \exp(xt) f_\xi(x) dx, \quad \alpha_s = \left. \frac{\partial^s \alpha_\xi(t)}{\partial t^s} \right|_{t=0} = \int_{-\infty}^{\infty} x^s f_\xi(x) dx.$$

Для случая центральных моментов имеем:

$$\begin{aligned} \mu_\xi(t) &= M(\exp\{(\xi - m_\xi)t\}), \\ \mu_\xi(t) &= M\left(1 + (\xi - m_\xi)t + (\xi - m_\xi)^2 \frac{t^2}{2!} + \dots\right) = \\ &= 1 + \mu_2 \frac{t^2}{2!} + \mu_3 \frac{t^3}{3!} + \dots \end{aligned}$$

и

$$\mu_s = \left. \frac{\partial^s \mu_\xi(t)}{\partial t^s} \right|_{t=0} = \int_{-\infty}^{\infty} (x - m_\xi)^s f_\xi(x) dx.$$

1.4.3 Семиинварианты или кумулянты

Пусть ξ – случайная величина с производящей функцией моментов $\alpha_\xi(t)$. Прологарифмируем $\alpha_\xi(t)$ и, после разложения в ряд, получаем

$$\log \alpha_\xi(t) = \varkappa_1 t + \varkappa_2 \frac{t^2}{2!} + \varkappa_3 \frac{t^3}{3!} + \dots,$$

где постоянные

$$\varkappa_s = \left. \frac{\partial^s \log \alpha_\xi(t)}{\partial t^s} \right|_{t=0}$$

называют *семиинвариантами* или *кумулянтами* ($\varkappa_1 = m_\xi$, $\varkappa_2 = D\xi$, $\varkappa_3 = \mu_3$, $\varkappa_4 = \mu_4 - 3\mu_2^2$).

1.5 Предельные теоремы

1.5.1 Сходимость по вероятности

Рассмотрим последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n \dots$. Вероятность

$$P(|\xi_n - a| < \varepsilon) \rightarrow 1 \quad \text{при } n \rightarrow \infty,$$

того, что $|\xi_n - a|$ меньше любого заданного числа $\varepsilon > 0$, то говорят, что последовательность ξ_n сходится к a по вероятности.

1.5.2 Неравенство Чебышева

Предположим, что ξ – произвольная случайная величина с плотностью вероятностей $f_\xi(x)$, с конечной дисперсией $D\xi$ и математическим ожиданием m_ξ , тогда выполняется неравенство

$$P(|\xi - m_\xi| \geq \alpha) \leq D\xi/\alpha^2.$$

Справедливо равенство

$$P(|\xi - m_\xi| \geq \alpha) = \int_{|\xi - m_\xi| \geq \alpha} f_\xi(x) dx.$$

Теперь мы можем записать дисперсию $D\xi$ в виде

$$\begin{aligned} D\xi &= \int_{-\infty}^{\infty} (x - m_\xi)^2 f_\xi(x) dx = \int_{-\infty}^{\infty} |x - m_\xi|^2 f_\xi(x) dx \geq \\ &\geq \int_{|x - m_\xi| \geq \alpha} |x - m_\xi|^2 f_\xi(x) dx \geq \alpha^2 \int_{|x - m_\xi| \geq \alpha} f_\xi(x) dx = \alpha^2 P(|\xi - m_\xi| \geq \alpha). \end{aligned}$$

1.5.3 Теорема Чебышева

Пусть ξ – случайная величина с математическим ожиданием m_ξ и дисперсией $D\xi$. Рассмотрим n реализаций случайной величины – $\xi_1, \xi_2, \dots, \xi_n$. Введем случайную величину $\eta = \sum_{i=1}^n \xi_i/n$ (среднее арифметическое)

$$m_\eta = M\eta = \frac{1}{n} \sum_{i=1}^n M\xi_i = \frac{1}{n} n m_\xi = m_\xi$$

и дисперсию

$$D\eta = \frac{1}{n^2} \sum_{i=1}^n D\xi_i = \frac{D\xi}{n}.$$

Мы покажем, что η сходится к m_ξ по вероятности:

$$P\left(\left|\frac{\sum_{i=1}^n \xi_i}{n} - m_\xi\right| < \varepsilon\right) > 1 - \delta.$$

Доказательство: Запишем неравенство Чебышева при $\alpha = \varepsilon$

$$P(|\eta - m_\eta| \geq \varepsilon) \leq \frac{D\eta}{\varepsilon^2} = \frac{D\xi}{n\varepsilon^2}.$$

Введя δ

$$\frac{D\xi}{n\varepsilon^2} < \delta,$$

получаем

$$P\left(\left|\frac{\sum_{i=1}^n \xi_i}{n} - m_\xi\right| \geq \varepsilon\right) \leq \delta$$

или

$$P\left(\left|\frac{\sum_{i=1}^n \xi_i}{n} - m_\xi\right| < \varepsilon\right) > 1 - \delta.$$

1.5.4 Центральная предельная теорема

Рассмотрим последовательность $\xi_1, \xi_2, \dots, \xi_n$ независимых случайных величин с плотностями вероятностей $f_\xi(x)$, математическим ожиданием m_ξ и дисперсией σ_ξ^2 , тогда $\eta_n = \sum_{i=1}^n \xi_i$ имеет нормальное распределение.

Доказательство: Предполагаем, что случайная величина ξ_i является непрерывной. Характеристическая функция случайной величины η :

$$g_\eta(t) = [g_\xi(t)]^n, \quad \text{where} \quad g_\xi(t) = \int_{-\infty}^{\infty} \exp(itx) f_\xi(x) dx.$$

Раскладывая в ряд $g_\xi(t)$, получаем

$$g_\xi(t) \approx g_\xi(0) + g'_\xi(0)t + [g''_\xi(0)/2 + \alpha(t)]t^2,$$

где $\alpha(t) \rightarrow 0$ для $t \rightarrow 0$, $g_\xi(0) = \int_{-\infty}^{\infty} f_\xi(x) dx = 1$,

$$g'_\xi(0) = \left[\int_{-\infty}^{\infty} ix \exp(itx) f_\xi(x) dx \right] \Big|_{t=0} = i \int_{-\infty}^{\infty} x f_\xi(x) dx = im_\xi.$$

Предположим, что $m_\xi = 0$, тогда $g'_\xi(0) = 0$, и

$$g''_\xi(0) = - \int_{-\infty}^{\infty} x^2 f_\xi(x) dx = -\sigma^2,$$

$$g_\xi(t) = 1 - \left[\frac{\sigma_\xi^2}{2} - \alpha(t) \right] t^2.$$

Вводя $\zeta_n = \eta_n / \sigma \sqrt{n}$, получаем

$$g_\zeta(t) = g_\eta\left(\frac{t}{\sigma \sqrt{n}}\right) = \left[g_\xi\left(\frac{t}{\sigma \sqrt{n}}\right) \right]^n$$

и

$$g_\zeta(t) = \left\{ 1 - \left[\frac{\sigma_\xi^2}{2} - \alpha\left(\frac{t}{\sigma \sqrt{n}}\right) \right] \frac{t^2}{\sigma_\xi n^2} \right\}^n.$$

Обозначая

$$\beta = \left[\frac{\sigma_\xi^2}{2} - \alpha \left(\frac{t}{\sigma_\xi \sqrt{n}} \right) \right] \frac{t^2}{\sigma_\xi^2 n}$$

и вычисляя логарифм функции $g_\zeta(t)$, получаем

$$\ln g_\zeta(t) = n \ln(1 - \beta).$$

Раскладывая в ряд

$$\ln(1 - \beta) = - \left[\beta + \frac{\beta^2}{2!} + \frac{\beta^3}{3!} + \dots \right]$$

при $n \rightarrow \infty$, можем записать

$$\ln(1 - \beta) \approx -\beta.$$

Таким образом, получаем

$$\lim_{n \rightarrow \infty} \ln g_{\zeta_n}(t) = \lim_{n \rightarrow \infty} n(-\beta) = \lim_{n \rightarrow \infty} \left\{ -\frac{t^2}{2} + \alpha \left(\frac{t}{\sigma_\xi \sqrt{n}} \right) \frac{t^2}{\sigma_\xi^2} \right\}.$$

По определению $\alpha(t) \rightarrow 0$ for $t \rightarrow 0$, следовательно:

$$\lim_{n \rightarrow \infty} \alpha \left(\frac{t}{\sigma_\xi \sqrt{n}} \right) = 0,$$

и

$$\lim_{n \rightarrow \infty} \ln g_{\zeta_n}(t) = -\frac{t^2}{2},$$

отсюда

$$\lim_{n \rightarrow \infty} g_{\zeta_n}(t) = \exp \left(-\frac{t^2}{2} \right)$$

и

$$f_\xi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left(-\frac{t^2}{2} \right) \exp(-itx) dt = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2} \right).$$

Таким образом получаем нормальное распределение с математическим ожиданием $m_\zeta = 0$ и стандартным отклонением $\sigma_\zeta = 1$.

1.6 Функции распределения дискретных случайных величин

Пусть ξ – дискретная случайная величина, которая может принимать любое из конечного числа значений x_1, x_2, \dots, x_n с вероятностями p_1, p_2, \dots, p_n ($\sum_{i=1}^n p_i = 1$, $P(\xi = x_i) = p_i$) (см. рис. 1.28 и табл. 1.1).

ξ	x_1	x_2	\cdots	x_n
P	p_1	p_2	\cdots	p_n

Таблица 1.1: Табличное представление вероятностей дискретной случайной величины.

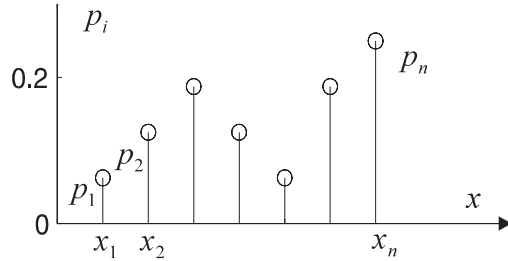


Рис. 1.28: Графическое представление вероятностей дискретной случайной величины.

1.6.1 Биномиальное распределение

Рассмотрим эксперимент с двумя возможными исходами, назове эти исходы благоприятным и неблагоприятным. Пусть p – вероятность благоприятного исхода. Если такой эксперимент повторяется n раз, то мы имеем дело с *биномиальным распределением*, которое дает вероятность реализации m благоприятных исходов в серии из n экспериментов.

Определение: Случайная величина ξ имеет *биномиальное распределение*, если ее распределение имеет вид

$$P_m = P(\xi = m) = C_n^m p^m q^{n-m}, \quad 0 < p < 1, \quad q = 1 - p, \quad m \in \{0, 1, \dots, n\}$$

(см. рис. 1.29).

Первые два момента распределения:

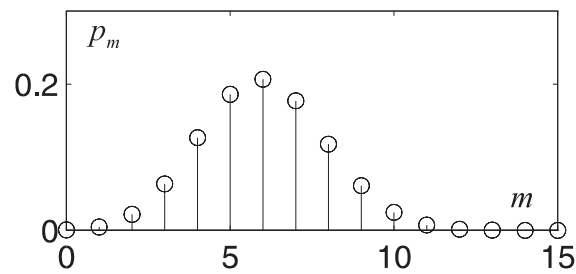
$$M\xi = np, \quad D\xi = np(1 - p).$$

Коэффициент асимметрии и эксцесс:

$$\gamma_1 = \frac{1 - 2p}{[np(1 - p)]^{1/2}} \quad \text{and} \quad \gamma_2 = \frac{1 - 6p(1 - p)}{np(1 - p)}.$$

Пример: Пусть вероятность появления события A равна p . Как много экспериментов n необходимо произвести, чтобы вероятность появления одного события была равна α

$$\begin{aligned} P(\xi \geq 1) &\geq \alpha, \\ 1 - P(\xi = 0) &\geq \alpha, \quad P(\xi = 0) \leq 1 - \alpha, \\ m = 0, \quad (1 - p)^n &\leq 1 - \alpha, \quad n \geq \ln(1 - \alpha) / \ln(1 - p). \end{aligned}$$

Рис. 1.29: Биномиальное распределение ($p = 0.4$, $n = 15$).

1.6.2 Распределение Пуассона

Дискретная случайная величина $\xi \{0, 1, \dots, m, \dots\}$ имеет *распределение Пуассона* с вещественным параметром a , если вероятность события $\xi = m$ описывается формулой

$$p_m = P(\xi = m) = \frac{a^m}{m!} e^{-a}, \quad a > 0$$

(см. рис. 1.30). Получим распределение Пуассона из биномиального распределения. Представим p в виде $p = a/n$, при этом полагаем, что $p \rightarrow 0$ при $n \rightarrow \infty$. Численные характеристики распределения Пуассона:

$$M\xi = a, \quad D\xi = a, \quad \gamma_1 = 1/\sqrt{a}, \quad \gamma_2 = 1/a.$$

Если случайные события являются пуассоновским потоком, то число событий ξ попадающих во временной интервал $(t_0, t_0 + \tau)$ имеет распределение Пуассона

$$p_m = \frac{a^m}{m!} e^{-a},$$

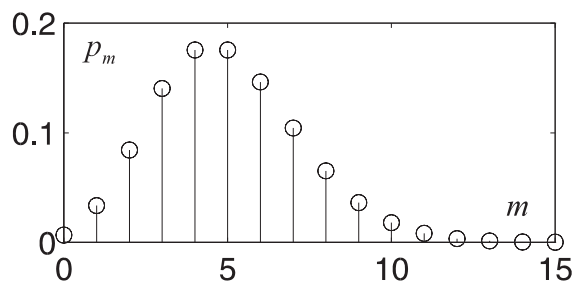
где a – математическое ожидание числа событий. Если $\lambda(t)$ – интенсивность потока, тогда a может быть представлена следующим образом

$$a = \int_{t_0}^{t_0 + \tau} \lambda(t) dt.$$

Модель пуассоновского потока хорошо описывает времена вступления сейсмических волн на сейсмограмме.

Пример: Представим, что зарегистрированы сейсмические волны отраженные от слоистой среды. Вступления волн описываются стационарным пуассоновским потоком с интенсивностью λ . Рассмотрим три ситуации, которые могут наблюдаться с течением времени τ :

- не будет зарегистрировано сейсмических волн (событие A);
- будет зарегистрировано не меньше трех волн (событие B);

Рис. 1.30: Распределение Пуассона ($a = 5$).

с) будут зарегистрированы три волны (событие C).

$$P(A) = p_0 = e^{-\lambda\tau}, \quad P(C) = \frac{(\lambda\tau)^3}{3!} e^{-\lambda\tau},$$

$$P(B) = 1 - (p_0 + p_1 + p_2) = 1 - e^{-\lambda\tau}[1 + \lambda\tau + 0.5(\lambda\tau)^2].$$

1.6.3 Геометрическое распределение

Дискретная случайная величина $\xi \{0, 1, 2, \dots, m, \dots\}$ имеет *геометрическое распределение*, если вероятность $\xi = m$ дается формулой

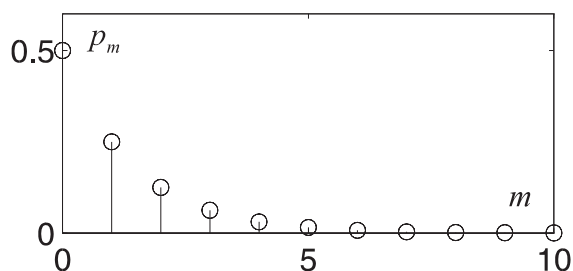
$$p_m = P(\xi = m) = q^m p, \quad 0 < p < 1, \quad q = 1 - p,$$

где p – вероятность благоприятного исхода q – вероятность неблагоприятного исхода (см. рис. 1.31).

Геометрическое распределение описывает вероятность реализации m неблагоприятных исходов перед появлением первого благоприятного исхода.

Численные характеристики геометрического распределения:

$$M\xi = (1-p)/p, \quad D\xi = (1-p)/p^2, \quad \gamma_1 = (2-p)/(1-p)^{1/2}, \quad \gamma_2 = (p^2 - 6p + 6)/(1-p).$$

Рис. 1.31: Геометрическое распределение ($p = 0.5$).

1.7 Непрерывные случайные величины

Пусть ξ – непрерывная случайная величина. Вероятность события $\xi < x$ описывается функцией распределения $F_\xi(x)$ и плотностью вероятностей $f_\xi(x)$

$$P(\xi < x) = F_\xi(x) = \int_{-\infty}^x f_\xi(t)dt, \quad f_\xi(x) \geq 0, \quad \int_{-\infty}^{\infty} f_\xi(x)dx = 1.$$

Вероятность события $\alpha < \xi \leq \beta$ можно выразить через функцию распределения

$$P(\alpha \leq \xi < \beta) = F_\xi(\beta) - F_\xi(\alpha).$$

1.7.1 Одномерное нормальное распределение

Одномерное нормальное распределение – наиболее важное распределение непрерывной случайной величины.

Определение: Случайная величина ξ имеет нормальное распределение с параметрами m_ξ и σ_ξ^2 (символическая запись – $\xi \in N(m_\xi, \sigma_\xi^2)$) если ее плотность вероятностей $f_\xi(x)$ имеет вид

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma_\xi} \exp\left\{-\frac{(x - m_\xi)^2}{2\sigma_\xi^2}\right\} \quad \text{for} \quad -\infty < x < \infty,$$

при этом функция распределения записывается следующим образом

$$F_\xi(x) = \Phi\left(\frac{x - m_\xi}{\sigma_\xi}\right),$$

где

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left\{-\frac{1}{2}y^2\right\}dy$$

– функция Лапласа (см. рис. 1.32).

Свойства функции Лапласа.

1. $\Phi(-\infty) = (1/\sqrt{2\pi}) \int_{-\infty}^{-\infty} \exp\{-\frac{1}{2}y^2\}dy = 0$.
2. $\Phi(-x) = 1 - \Phi(x)$.
3. $\Phi(\infty) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}y^2\}dy = 1$.

Характеристическая функция:

$$g(t) = \exp\{itm_\xi - t^2\sigma_\xi^2/2\}.$$

Коэффициент асимметрии и эксцесс равны нулю – $\gamma_1 = \gamma_2 = 0$. Моменты:

$$\mu_{2s} = \frac{(2s)!}{2^s s!} \sigma_\xi^{2s}, \quad \mu_{2s+1} = 0, \quad s \geq 1.$$

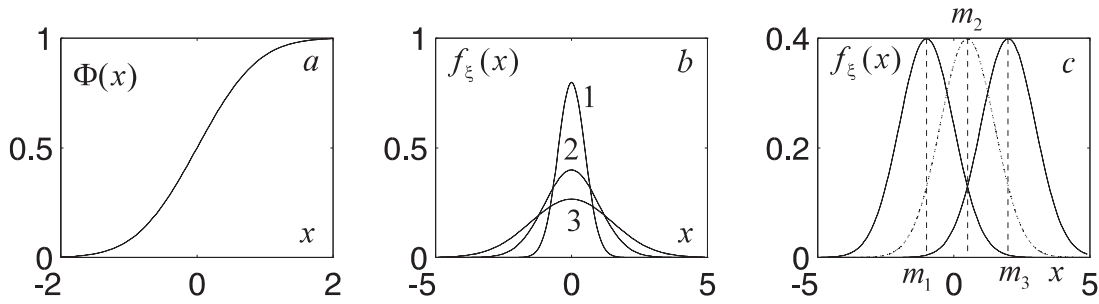


Рис. 1.32: Функция Лапласа и нормальная плотность вероятностей. (а) – функция Лапласа: $m = 0$, $\sigma = 0.5$. (б) – нормальная функция распределения: $m_1 = m_2 = m_3 = 0$; $\sigma_1 = 0.5$ – (1); $\sigma_2 = 1.0$ – (2); $\sigma_3 = 1.5$ – (3). (с) нормальная функция распределения: $m_1 = -1$; $m_2 = 0.5$; $m_3 = 2$; $\sigma_1 = \sigma_2 = \sigma_3 = 1$.

Example: Для случайной величины ξ с $\xi \in N(m_\xi, \sigma_\xi^2)$ вероятности $P = 0.9$, $P = 0.95$ and $P = 0.99$ могут быть получены численно:

$$P(-1.64 < \frac{\xi - m_\xi}{\sigma_\xi} < 1.64) = 0.90,$$

$$P(-1.96 < \frac{\xi - m_\xi}{\sigma_\xi} < 1.96) = 0.95,$$

$$P(-2.58 < \frac{\xi - m_\xi}{\sigma_\xi} < 2.58) = 0.99.$$

Вводят в рассмотрение *стандартную нормальную плотность* ($N(0, 1)$):

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}t^2\right\},$$

стандартную нормальную функцию распределения ($\Phi(x)$):

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{1}{2}t^2\right\} dt.$$

Если случайные величины $\xi_1, \xi_2, \dots, \xi_n$ имеют нормальное распределение и независимы, тогда произвольная линейная комбинация $\eta = \sum_{i=1}^n a_i \xi_i$ of ξ_i также имеет нормальное распределение.

Предположим, что ξ_1 и ξ_2 – независимы и распределены по нормальному закону с математическими ожиданиями m_1 и m_2 , дисперсиями σ_1^2 и σ_2^2 , тогда $a_1 \xi_1$ и $a_2 \xi_2$ нормально распределены с характеристическими функциями

$$g_{a_1 \xi_1}(t) = \exp(it a_1 m_1 - t^2 a_1^2 \sigma_1^2 / 2), \quad g_{a_2 \xi_2}(t) = \exp(it a_2 m_2 - t^2 a_2^2 \sigma_2^2 / 2).$$

Характеристическая функция случайной величины $\eta = a_1 \xi_1 + a_2 \xi_2$ равна

$$g_\eta(t) = g_{a_1 \xi_1}(t) g_{a_2 \xi_2}(t) = \exp[it(a_1 m_1 + a_2 m_2) - t^2(a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2) / 2],$$

т.е. случайная величина η принадлежит нормальному распределению с математическим ожиданием $m_\eta = a_1 m_1 + a_2 m_2$ и дисперсией $\sigma_\eta^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2$.

Аналогично можно показать, что

$$\eta = \sum_{i=1}^n a_i \xi_i$$

будет иметь нормальное распределение с математическим ожиданием

$$m_\eta = \sum_{i=1}^n a_i m_i$$

и дисперсией

$$\sigma_\eta^2 = \sum_{i=1}^n a_i^2 \sigma_i^2.$$

1.7.2 Многомерное нормальное распределение

Случайный вектор $\vec{\xi} = [\xi_1, \xi_2, \dots, \xi_n]^T$ имеет нормальное распределение (символическая запись – $N(\vec{m}_\xi, D_\xi)$) если плотность вероятностей $f_\xi(\vec{x})$ вектора \vec{x} дается формулой

$$f_\xi(\vec{x}) = (2\pi)^{-n/2} |D_\xi|^{-1/2} \exp \left\{ -(1/2)(\vec{x} - \vec{m}_\xi)^T D_\xi^{-1} (\vec{x} - \vec{m}_\xi) \right\},$$

где \vec{m}_ξ – вектор математического ожидания вектора $\vec{\xi}$: $M\vec{\xi} = \vec{m}_\xi$, D_ξ – ковариационная матрица вектора $\vec{\xi}$: $[D_\xi]_{ij} = M[(\xi_i - m_{\xi_i})(\xi_j - m_{\xi_j})]$. Характеристическая функция дается выражением:

$$g_\xi(\vec{t}) = \exp[i\vec{t}^T \vec{m}_\xi - (1/2)\vec{t}^T D_\xi \vec{t}].$$

Пример: Две случайные величины ξ_1 and ξ_2 имеют нормальное распределение если их плотность вероятностей можно записать в виде

$$f_{\xi_1 \xi_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp \left\{ -\frac{1}{2(1-r^2)} \left[\frac{(x_1 - m_1)^2}{\sigma_1^2} - 2r \frac{(x_1 - m_1)(x_2 - m_2)}{\sigma_1\sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2} \right] \right\},$$

(см. рис. 1.33) с ковариационной матрицей

$$D = \begin{vmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix}, \quad r = \frac{M[(\xi_1 - m_1)(\xi_2 - m_2)]}{\sigma_1\sigma_2}.$$

Маргинальные плотности ξ_1 и ξ_2 – принадлежат нормальному распределению и даются формулами:

$$f_{\xi_1}(x_1) = \int_{-\infty}^{\infty} f_{\xi_1 \xi_2}(x_1, x_2) dx_2, \quad f_{\xi_2}(x_2) = \int_{-\infty}^{\infty} f_{\xi_1 \xi_2}(x_1, x_2) dx_1,$$

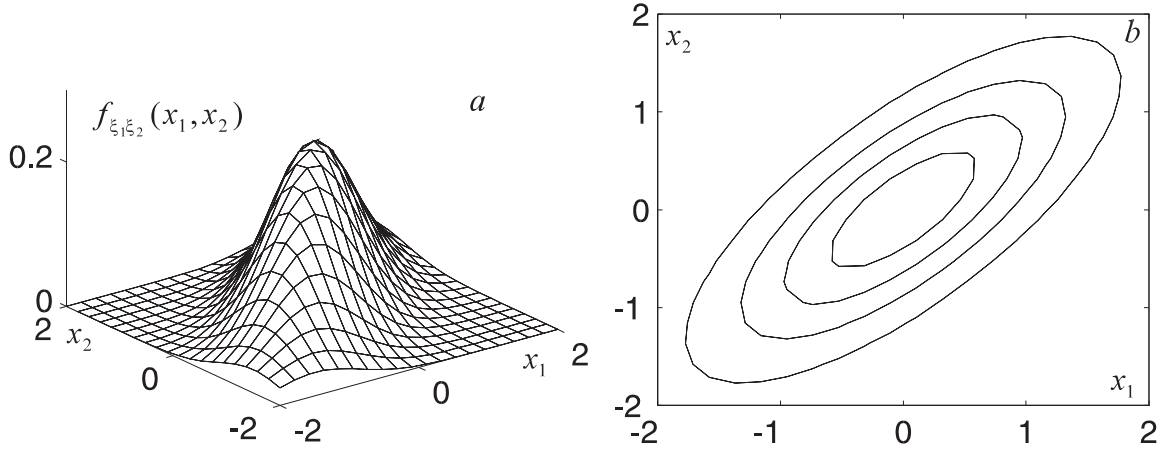


Рис. 1.33: Нормальная плотностей вероятностей двух случайных величин ξ_1 и ξ_2 : $m_1 = m_2 = 0$; $\sigma_1 = \sigma_2 = 1$; $r = 0.75$. (а) – трехмерное изображение; (б) – линии постоянных значений плотности вероятностей на плоскости (x_1, x_2) .

$$f_{\xi_1}(x_1) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left\{-\frac{(x_1 - m_1)^2}{2\sigma_1^2}\right\},$$

$$f_{\xi_2}(x_2) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left\{-\frac{(x_2 - m_2)^2}{2\sigma_2^2}\right\}.$$

Условная плотность ξ_2 , в предположении $\xi_1 = x_1$ записывается в виде

$$\begin{aligned} f_{\xi_2/\xi_1}(x_2/x_1) &= \frac{f_{\xi_1, \xi_2}(x_1, x_2)}{f_{\xi_1}(x_1)} = \\ &= \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} \exp\left\{\frac{-1}{2(1-r^2)} \left[\frac{(x_2 - m_2)}{\sigma_2} - r\frac{(x_1 - m_1)}{\sigma_1}\right]^2\right\} = \\ &= \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} \exp\left\{\frac{-1}{2(1-r^2)\sigma_2^2} \left[(x_2 - m_2) - r\frac{\sigma_2}{\sigma_1}(x_1 - m_1)\right]^2\right\}. \end{aligned}$$

Условная плотность ξ_1 , в предположении $\xi_2 = x_2$ дается следующей формулой

$$f_{\xi_1/\xi_2}(x_1/x_2) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-r^2)\sigma_1^2} \left[(x_1 - m_1) - r\frac{\sigma_1}{\sigma_2}(x_2 - m_2)\right]^2\right\}.$$

Условные плотности вероятностей являются нормальными с математическим ожиданием и дисперсией:

$$M[\xi_2/\xi_1 = x_1] = m_{\xi_2/\xi_1} = m_2 + r\frac{\sigma_2}{\sigma_1}(x_1 - m_1), \quad \sigma_{\xi_2/\xi_1} = \sigma_2\sqrt{1-r^2},$$

$$M[\xi_1/\xi_2 = x_2] = m_{\xi_1/\xi_2} = m_1 + r\frac{\sigma_1}{\sigma_2}(x_2 - m_2), \quad \sigma_{\xi_1/\xi_2} = \sigma_1\sqrt{1-r^2}$$

(см. рис. 1.34).

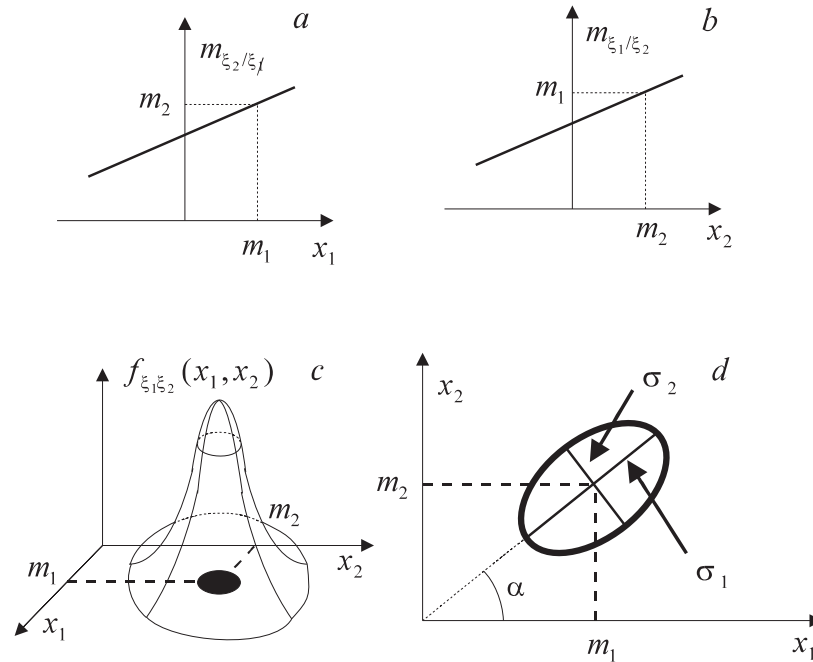


Рис. 1.34: Двумерное нормальное распределение. (а) и (b) – условные математические ожидания ξ_1 и ξ_2 соответственно; (с) – графическое представление плотности вероятностей; (d) – контуры плотности вероятностей на плоскости (x_1, x_2) .

$$\frac{(x_1 - m_1)^2}{\sigma_1^2} - 2r \frac{(x_1 - m_1)(x_2 - m_2)}{\sigma_1 \sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2} = \lambda^2,$$

$$\tan 2\alpha = \frac{2r\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}.$$

1.7.3 Равномерное распределение

Если плотность вероятностей случайной величины ξ имеет прямоугольный вид

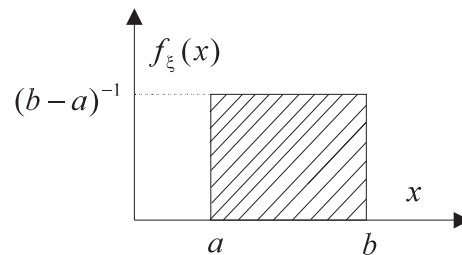


Рис. 1.35: Равномерное распределение.

$$f_\xi(x) = \begin{cases} (b-a)^{-1} & x \in (a, b), \\ 0 & x \notin (a, b), \end{cases}$$

то ξ распределена по равномерному закону на интервале (a, b) (см. рис. 1.35)
Числовые характеристики распределения:

$$M\xi = (a + b)/2, \quad D\xi = (b - a)^2/12, \quad \gamma_1 = 0, \quad \gamma_2 = -1.2.$$

1.7.4 Распределение χ^2

Рассмотрим случайный вектор $\vec{\xi} = [\xi_1, \dots, \xi_n]^T$ размерности $n \times 1$ с нормальным распределением $\vec{\xi} = N(0, 1)$. Сумма квадратов компонент этого вектора $\eta = \sum_{i=1}^n \xi_i^2$ распределена по закону χ_n^2 с n степенями свободы

$$\eta = \sum_{i=1}^n \xi_i^2, \quad \xi_i \in N(0, 1), \quad \eta \in \chi_n^2,$$

Соответствующая плотность вероятностей дается формулой

$$f_\eta(x) = \begin{cases} \left[(x/2)^{(n/2)-1} \exp(-x/2) \right] / (2\Gamma(n/2)) & x > 0, \\ 0 & x \leq 0 \end{cases}$$

(см. рис. 1.36).

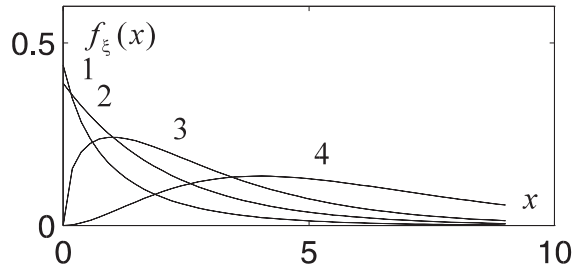


Рис. 1.36: Распределение χ^2 . 1 - $n = 1$; 2 - $n = 2$; 3 - $n = 3$; 4 - $n = 6$.

Моменты распределения χ_n^2 :

$$M\xi = n, \quad D\xi = 2n, \quad \gamma_1 = 2\sqrt{2/n}, \quad \gamma_2 = 12/n.$$

Характеристическая функция:

$$g(t) = (1 - 2it)^{-n/2}.$$

Пусть $\vec{\xi} = [\xi_1, \dots, \xi_n]^T$ — случайный вектор с компонентами, распределенными по нормальному закону $\xi_i \in N(m, \sigma^2)$. Выборочное среднее:

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

Выборочная дисперсия:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \bar{\xi})^2.$$

Случайная величина

$$\eta = \frac{(n-1)s^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (\xi_i - \bar{\xi})^2 \in \chi_{n-1}^2$$

распределена по закону χ^2 .

1.7.5 Распределение Стьюдента (t-распределение)

Пусть случайные переменные α и ζ независимы и имеют распределения $\alpha \in N(0, 1)$ и $\zeta \in \chi_n^2$ соответственно, тогда случайная величина

$$\eta = \frac{\alpha}{(\zeta/n)^{1/2}}$$

принадлежит *t-распределению* (другое название – *распределение Стьюдента*) с n степенями свободы и плотностью

$$f_\eta(t) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}$$

(см. рис. 1.37).

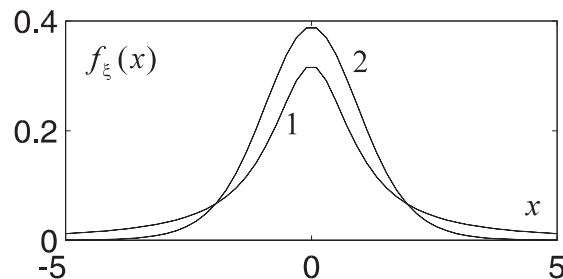


Рис. 1.37: Распределение Стьюдента. 1 — $n = 1$; 2 — $n = 10$.

Пусть ξ_1, \dots, ξ_n — случайный вектор, принадлежащий нормальному распределению $\xi_i \in N(m, \sigma^2)$.

Выборочное математическое ожидание:

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

Выборочная дисперсия:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \bar{\xi})^2.$$

Запишем:

$$\eta = \sqrt{n}(\bar{\xi} - m_\xi)/s.$$

Моменты:

$$\begin{aligned} M\eta &= 0, & D\eta &= n/(n-2), & n &> 2, \\ \gamma_1 &= 0, & \gamma_2 &= 6/(n-4), & n &> 4. \end{aligned}$$

1.7.6 Распределение Фишера

Пусть случайные величины $\xi_1, \xi_2, \dots, \xi_{n_1}$ и $\eta_1, \eta_2, \dots, \eta_{n_2}$ независимы и принадлежат распределениям:

$$\xi_i \in N(m_1, \sigma_1^2), \quad \eta_i \in N(m_2, \sigma_2^2).$$

введем в рассмотрение независимые случайные величины s_1 и s_2

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (\xi_i - \bar{\xi})^2, \quad s_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (\eta_i - \bar{\eta})^2.$$

Независимые случайные величины ζ_1 and ζ_2 принадлежат соответственно распределениям

$$\zeta_1 = \frac{n_1 s_1^2}{\sigma_1^2} \in \chi_{n_1}^2, \quad \zeta_2 = \frac{n_2 s_2^2}{\sigma_2^2} \in \chi_{n_2}^2,$$

при этом случайная величина

$$\zeta = \frac{\zeta_1/n_1}{\zeta_2/n_2} = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \in F(n_1, n_2)$$

принадлежит распределению Фишера $F(n_1, n_2)$ с n_1 и n_2 степенями свободы, с плотностью вероятностей

$$f_\zeta(F) = \frac{n_1^{n_2/2} n_2^{n_1/2} F^{(n_1+n_2)/2}}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} \frac{F^{(n_1/2)-1}}{(n_1 + n_2 F)^{(n_1+n_2)/2}}$$

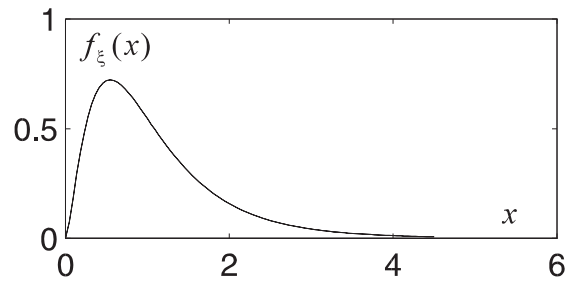
(см. рис. 1.38).

Моменты распределения Фишера:

$$M\zeta = \frac{n_2}{n_2 - 2}, \quad n_2 > 2; \quad D\zeta = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}, \quad n_2 > 4.$$

Плотность вероятностей $f_\zeta(F)$ близка к нормальной плотности при достаточно больших значениях n_1 и n_2 ($n_1, n_2 > 50$)

$$f_\zeta(F) \xrightarrow{n_1, n_2 \rightarrow \infty} N(m, \sigma).$$

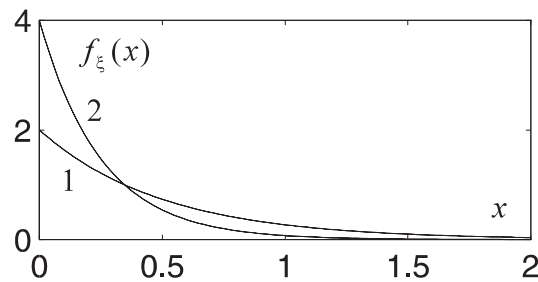
Рис. 1.38: Распределение Фишера ($n_1 = 5$, $n_2 = 50$).

1.7.7 Экспоненциальное распределение

Плотность вероятностей экспоненциального распределения имеет вид

$$f_{\xi}(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{for } x \geq 0, \\ 0 & \text{for } x < 0 \end{cases}$$

где $\lambda > 0$ (см. рис. 1.39).

Рис. 1.39: Экспоненциальное распределение. 1 – $\lambda = 2$, 2 – $\lambda = 4$.

Моменты:

$$M\xi = \lambda^{-1}, \quad D\xi = \lambda^{-2}, \quad \gamma_1 = 2, \quad \gamma_2 = 6.$$

Характеристическая функция:

$$g(t) = (1 - it/\lambda)^{-1}.$$

Экспоненциальное распределение играет большую роль в теории марковских процессов. В случае стационарного пуассоновского потока с интенсивностью λ , случайный временной интервал между соседними событиями описывается экспоненциальным распределением с параметром λ .

1.7.8 Распределение Лапласа

Случайная величина ξ принадлежит распределению Лапласа, если ее плотность вероятностей дается формулой

$$f_{\xi}(x) = (\lambda/2) \exp(-\lambda|x|)$$

(см. рис. 1.40).

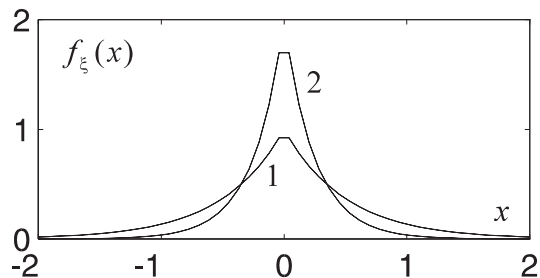


Рис. 1.40: Распределение Лапласа. 1 – $\lambda = 2$, 2 – $\lambda = 4$.

Моменты:

$$M\xi = 0, \quad D\xi = 2/\lambda^2.$$

1.7.9 Распределение Коши

Случайная величина ζ имеет распределение Коши, если ее плотность вероятностей дается формулой

$$f_\zeta(x) = \frac{1}{\pi} \frac{1}{1+x^2}$$

(см. рис. 1.41).

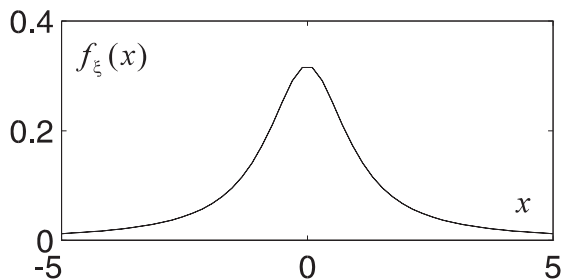


Рис. 1.41: Распределение Коши.

Распределению Коши принадлежит отношение двух случайных величин, распределенных по нормальному закону:

$$\zeta = \xi/\eta, \quad \xi \in N(0, 1), \quad \eta \in N(0, 1).$$

Характеристическая функция: $g(t) = \exp(-|t|)$.

Математическое ожидание $M\xi$ и другие моменты не определены, поскольку соответствующие интегралы расходятся.

1.7.10 Логарифмически нормальное распределение

Если случайная величина $\eta = \ln \xi$ принадлежит нормальному распределению, то случайная величина ξ имеет логарифмически нормальное распределение

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{\frac{-1}{2\sigma^2}(\ln x - m)^2\right\}$$

(см. рис. 1.42).

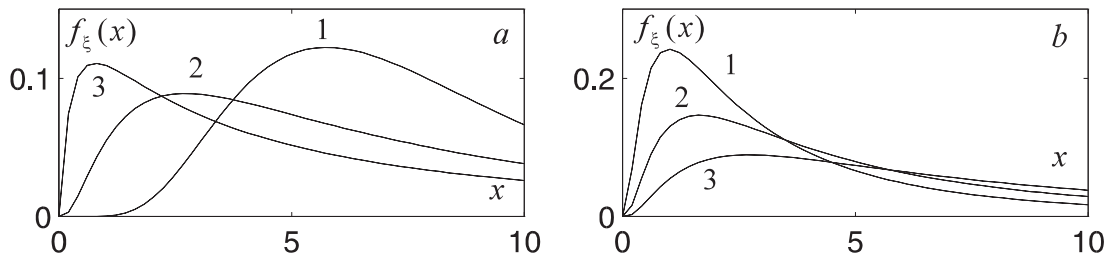


Рис. 1.42: Логарифмически нормальное распределение. (a) $m = 2.0$, $1 - \sigma = 0.5$, $2 - \sigma = 1.0$, $3 - \sigma = 1.5$; (b) $\sigma = 1.0$, $1 - m = 1.0$, $2 - m = 1.5$, $3 - m = 2.0$.

Моменты:

$$M\xi = \exp(m - \sigma^2/2), \quad D\xi = \exp(2m + \sigma^2)(\exp(\sigma^2) - 1).$$

1.7.11 Значение нормального распределения

Распределение суммы n независимых произвольно распределенных случайных величин стремится к нормальному распределению при стремлении n к бесконечности. Взаимосвязь между нормальным распределением и распределениями, рассмотренными ранее представлена на рис. 1.43 и рис. 1.44.

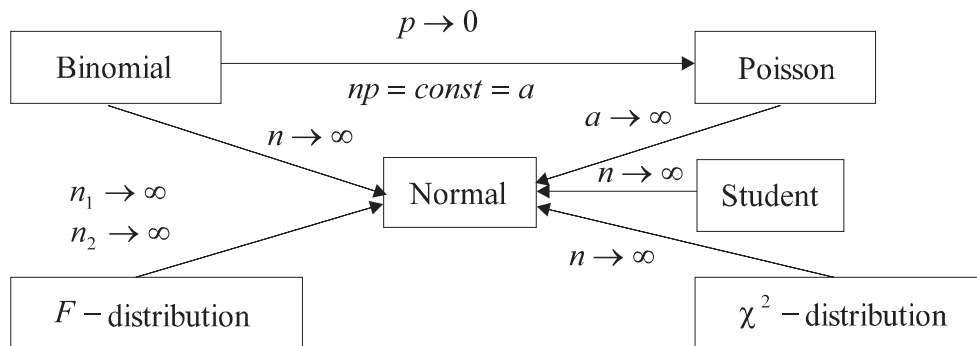


Рис. 1.43: Связь нормального распределения с некоторыми распределениями при предельных значениях их параметров.

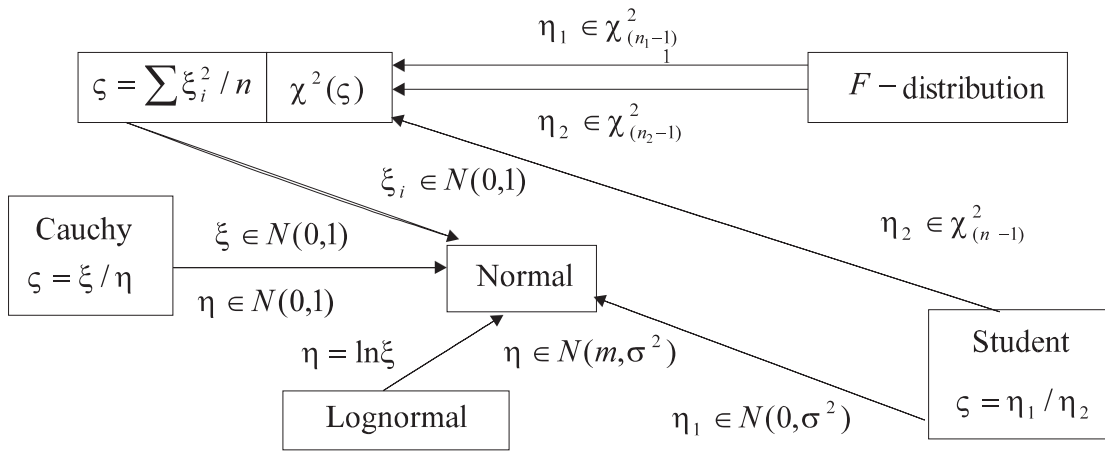


Рис. 1.44: Схема связи нормального распределения с некоторыми другими распределениями с точки зрения соотношений между соответствующими им случайных величин.

1.7.12 Доверительные интервалы

Пусть $\tilde{\rho}$ – оценка неизвестного параметра ρ . Доверительный интервал определяется следующим образом I_β

$$P(|\tilde{\rho} - \rho| < \varepsilon) = \beta, \quad P(\tilde{\rho} - \varepsilon < \rho < \tilde{\rho} + \varepsilon) = \beta, \quad I_\beta = (\tilde{\rho} - \varepsilon, \tilde{\rho} + \varepsilon),$$

где β – доверительная вероятность.

Доверительный интервал для математического ожидания

В качестве оценки математического ожидания будем рассматривать выборочное среднее

$$\tilde{\rho} = \bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i.$$

1. Рассмотрим случай известного среднеквадратического отклонения σ_ξ :

$$\begin{aligned} \xi_1, \xi_2, \dots, \xi_n; \quad m_\xi, \quad \sigma_m = \sigma_\xi / \sqrt{n}, \\ P\{-l \leq (\bar{\xi} - m_\xi) \sqrt{n} / \sigma_\xi \leq l\} = 2\Phi(l) - 1, \\ P\{\bar{\xi} - l\sigma_\xi / \sqrt{n} \leq m_\xi \leq \bar{\xi} + l\sigma_\xi / \sqrt{n}\} = 2\Phi(l) - 1. \end{aligned}$$

2. Рассмотрим случай неизвестного среднеквадратического отклонения σ_ξ :

$$\begin{aligned} \xi_1, \xi_2, \dots, \xi_n; \quad s^2 = 1/(n-1) \sum_{i=1}^n (\xi_i - \bar{\xi})^2, \\ \eta = \sqrt{n}(\bar{\xi} - m_\xi) / s, \quad P\{|\eta| < t_\beta\} = \beta. \end{aligned}$$

Случайная величина $\eta = \sqrt{n}(\bar{\xi} - m_\xi)/s$ принадлежит распределению Стьюдента (см. (1.7.5)) и по величине β можно найти величину t_β (см. рис. 1.45).

$$P\{\bar{\xi} - t_\beta s/\sqrt{n} \leq m_\xi \leq \bar{\xi} + t_\beta s/\sqrt{n}\} = \beta.$$

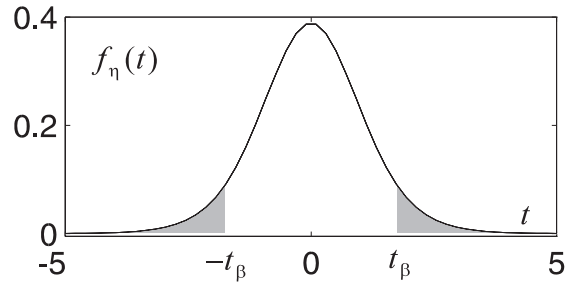


Рис. 1.45: Плотность вероятностей распределения Стьюдента ($n=10$) и 5.5 % точки.

Доверительный интервал для дисперсии

Пусть $\xi_1, \xi_2, \dots, \xi_n$ – случайная величина

$$\xi_i \in N(m_\xi, \sigma_\xi^2), \quad \frac{(n-1)s^2}{\sigma_\xi^2} \in \chi_{n-1}^2.$$

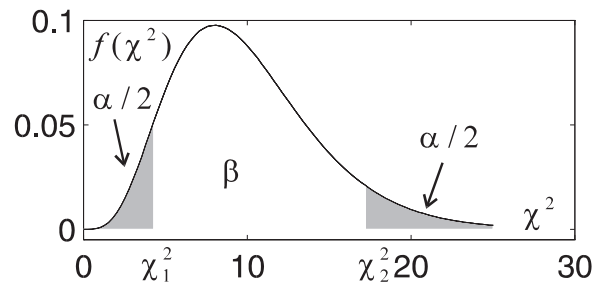


Рис. 1.46: Плотность вероятностей распределения χ^2 . $\alpha=0.13$

$$1 - P(\chi^2 > \chi_1^2) = P(\chi^2 < \chi_1^2) = \alpha/2,$$

$$P(\chi^2 > \chi_2^2) = \alpha/2.$$

Таким образом мы можем найти величины χ_1^2 и χ_2^2 (см. рис. 1.46)

$$P(\chi_1^2 < (n-1)s^2/\sigma^2 < \chi_2^2) = 1 - P(\chi^2 < \chi_1^2) - P(\chi^2 > \chi_2^2) = 1 - \alpha,$$

$$P\left(\frac{(n-1)s^2}{\chi_2^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_1^2}\right) = 1 - \alpha.$$

Глава 2

Информация и энтропия

Энтропия есть мера неопределенности состояния физической системы.

2.1 Энтропия множества дискретных состояний системы

Пусть x_1, x_2, \dots, x_n – возможные состояния системы с вероятностями p_1, p_2, \dots, p_n , тогда энтропия дается формулой

$$H = - \sum_{i=1}^n p_i \log p_i.$$

В случае равновероятных состояний $p_i = 1/n$, мы имеем:

$$H = -n \frac{1}{n} \log \frac{1}{n} = \log n.$$

Энтропию системы определяют следующим образом

$$H_\xi = M[-\log p_\xi].$$

2.2 Энтропия сложной системы

Пусть случайные величины ξ и η имеют соответственно реализации x_1, x_2, \dots, x_n и y_1, y_2, \dots, y_m (их совместное распределение может быть представлено в табличной форме – см. табл. 2.1). Энтропия (совместная) дается формулой

$$H_{\xi\eta} = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log p_{ij},$$
$$H_{\xi\eta} = M[-\log p(\xi, \eta)].$$

$y_j \backslash x_i$	x_1	x_2	\dots	x_{n-1}	x_n
y_1	p_{11}	p_{12}	\dots	$p_{1\ n-1}$	$p_{1\ n}$
y_2	p_{21}	p_{22}	\dots	$p_{2\ n-1}$	$p_{2\ n}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
y_m	p_{m1}	p_{m2}	\dots	$p_{m\ n-1}$	$p_{m\ n}$

Таблица 2.1: Табличное представление совместного распределения вероятностей двух дискретных случайных величин x_i and y_j .

Если случайные величины ξ и η независимы, то совместная вероятность записывается следующим образом

$$p(\xi, \eta) = p(\xi)p(\eta),$$

при этом энтропия имеет вид

$$H_{\xi\eta} = H_{\xi} + H_{\eta}.$$

В случае n случайных величин справедлива формула

$$H_{\xi_1, \xi_2, \dots, \xi_n} = H_{\xi_1} + H_{\xi_2} + \dots + H_{\xi_n}.$$

Пусть случайные величины ξ и η независимы, тогда условная энтропия системы ξ при условии, что случайная величина η принимает значение y_j :

$$H_{\xi/y_j} = - \sum_{i=1}^n p(x_i/y_j) \log p(x_i/y_j)$$

или

$$H_{\xi/y_j} = M[-\log p(\xi/y_j)].$$

Средняя или полная энтропия: The mean or perfect entropy of the system is

$$H_{\xi/\eta} = \sum_{j=1}^m p_j H_{\xi/y_j} = \sum_{i=1}^n \sum_{j=1}^m p_j p(x_i/y_j) \log p(x_i/y_j).$$

Принимая во внимание равенство

$$p_j p(x_i/y_j) = p_{ij},$$

получаем

$$H_{\xi/\eta} = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log p(x_i/y_j).$$

Энтропию системы (ξ, η) можно записать в виде

$$H_{\xi, \eta} = H_{\xi} + H_{\eta/\xi}$$

или

$$H_{\xi,\eta} = H_\eta + H_{\xi/\eta}.$$

Это легко доказывается, если воспользоваться определением энтропии и теоремой о произведении вероятностей

$$p(\xi, \eta) = p(\xi)p(\eta/\xi) = p(\eta)p(\xi/\eta).$$

2.3 Информация Шеннона (дискретный случай)

Количество получаемой информации связывают с уменьшением неопределенности системы

Пусть ξ – случайная величина, описывающая возможные состояния физической системы. Пусть энтропия системы до проведения наблюдений (априорная энтропия) – $H_\xi^{(apr)}$ и энтропия после проведения наблюдений (апостериорная энтропия) – $H_\xi^{(apost)}$, тогда информация Шеннона определяется следующим образом

$$I_\xi^{(s)} = H_\xi^{(apr)} - H_\xi^{(apost)}.$$

В частном случае, если результат наблюдения состояний системы является полностью определенным, то

$$H_\xi^{(apost)} = 0$$

и информация равна

$$I_\xi^{(s)} = H_\xi^{(apr)} = - \sum_{i=1}^n p_i \log p_i$$

или

$$I_\xi^{(s)} = M[-\log p(\xi)].$$

Количество информации о состоянии системы ξ , получаемое по наблюдениям системы η определяется выражением

$$I_{\eta \rightarrow \xi}^{(s)} = H_\xi - H_{\xi/\eta}.$$

Информация $I_{\eta \rightarrow \xi}^{(s)}$ называется средней или полной вероятностью, содержащейся в системе η относительно системы ξ . Легко показать, что

$$I_{\xi \rightarrow \eta}^{(s)} = H_\eta - H_{\eta/\xi},$$

Принимая во внимание

$$H_{\xi\eta} = H_\xi + H_{\eta/\xi} = H_\eta + H_{\xi/\eta},$$

получаем

$$I_{\xi \rightarrow \eta}^{(s)} = I_{\eta \rightarrow \xi}^{(s)} = I_{\eta \leftrightarrow \xi}^{(s)},$$

Информация $I_{\eta \leftrightarrow \xi}^{(s)}$ называется *полной взаимной информацией*.

Если случайные величины ξ и η независимы, то

$$I_{\eta \leftrightarrow \xi} = 0, \quad \text{as} \quad H_{\eta/\xi} = H_{\eta} \quad \text{and} \quad H_{\xi/\eta} = H_{\xi}.$$

Можно представить величину $I_{\eta \leftrightarrow \xi}$ через энтропию объединенной системы

$$I_{\eta \leftrightarrow \xi} = H_{\xi} + H_{\eta} - H_{\xi\eta}$$

и, используя оператор математического ожидания:

$$I_{\eta \leftrightarrow \xi} = M[-\log p(\xi) - \log p(\eta) + \log p(\xi, \eta)],$$

$$I_{\eta \leftrightarrow \xi} = M\left[\log \frac{p(\xi, \eta)}{p(\xi)p(\eta)}\right]$$

или

$$I_{\eta \leftrightarrow \xi} = \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log \frac{p_{ij}}{p_{\xi i} p_{\eta j}}.$$

Частная информация о системе ξ , содержащаяся в конкретном состоянии y_i обозначается как $I_{y_i \rightarrow \xi}$. Полная информация есть математическое ожидание частной информации

$$I_{\eta \leftrightarrow \xi} = \sum_{j=1}^m p_{\eta j} I_{y_j \rightarrow \xi},$$

где

$$I_{y_j \rightarrow \xi} = \sum_{i=1}^n p(x_i/y_j) \log \frac{p(x_i/y_j)}{p_{\xi i}}$$

или

$$I_{y_j \rightarrow \xi} = M_{y_i} \left[\log \frac{p(\xi/y_j)}{p(\xi)} \right],$$

где M_{y_i} – оператор условного математического ожидания. Частная информация о событии x_i , полученная при реализации события y_j , может быть представлена в виде

$$I_{y_j \rightarrow x_i} = \log \frac{p(x_i/y_j)}{p_{\xi i}}.$$

2.4 Непрерывный случай (энтропия и информация)

Энтропия системы описываемой непрерывной случайной величиной ξ с плотностью вероятностей $f(x)$, определяется как

$$H_{\xi} = - \int_{-\infty}^{\infty} f(x) \log f(x) dx$$

или

$$H_\xi = M[-\log f(x)].$$

В случае двух непрерывных систем ξ и η с плотностью вероятностей $f(x, y)$, определяют *частную условную энтропию*, т.е. энтропия системы η определяется тем, что система ξ имеет определенное состояние x :

$$H(\eta/x) = - \int_{-\infty}^{\infty} f(y/x) \log f(y/x) dy.$$

Полная или *средняя условная энтропия* определяется следующим образом

$$H(\eta/\xi) = - \int \int_{-\infty}^{\infty} f_\xi(x) f(y/x) \log f(y/x) dx dy$$

или, учитывая

$$f_{\xi\eta}(x, y) = f_\xi(x) f_{\eta/\xi}(y/x),$$

можно записать

$$H(\eta/\xi) = - \int \int_{-\infty}^{\infty} f_{\xi\eta}(x, y) \log f_{\eta/\xi}(y/x) dx dy.$$

По аналогии с дискретным случаем можно показать, что энтропия системы (ξ, η) равна

$$H(\xi, \eta) = H(\xi) + H(\eta/\xi).$$

Для независимых случайных величин ξ и η имеем

$$H(\xi, \eta) = H(\xi) + H(\eta).$$

Example: Найти плотность вероятностей $f_\xi(x)$ случайной величины ξ в классе непрерывных распределений, которая доставляет максимум энтропии H_ξ ($\max H_\xi$) при заданном математическом ожидании m_ξ и дисперсии σ_ξ^2 . Пусть

$$\int_{-\infty}^{\infty} f_\xi(x) dx = 1, \quad \int_{-\infty}^{\infty} x f_\xi(x) dx = m_\xi, \quad \int_{-\infty}^{\infty} (x - m_\xi)^2 f_\xi(x) dx = \sigma_\xi^2.$$

Найдем

$$\hat{f}_\xi(x) \Rightarrow \max \left[- \int_{-\infty}^{\infty} f_\xi(x) \log f_\xi(x) dx \right].$$

Это типичная вариационная проблема

$$\hat{f}_\xi(x) \Rightarrow \max \left[\int_{-\infty}^{\infty} \Phi(f_\xi(x), x) dx \right]$$

с дополнительными условиями

$$\int_{-\infty}^{\infty} \varphi_k(f_\xi(x), x) dx = c_k,$$

где

$$\begin{aligned} \Phi(f_\xi(x), x) &= -f_\xi(x) \log f_\xi(x), \\ \varphi_1 &= f_\xi(x), \quad c_1 = 1, \\ \varphi_2 &= x f_\xi(x), \quad c_2 = m_\xi, \\ \varphi_3 &= (x - m_\xi)^2 f_\xi(x), \quad c_3 = \sigma_\xi^2. \end{aligned}$$

Используя метод неопределенных множителей Лагранжа, запишем уравнение Эйлера

$$\partial \Phi_1 / \partial f = 0,$$

где

$$\Phi_1(f, x) = -f \log f + \lambda_1 f + \lambda_2 x f + \lambda_3 (x - m_\xi)^2 f.$$

После дифференцирования по f_1 получаем

$$-1 - \log f + \lambda_1 + \lambda_2 x + \lambda_3 (x - m_\xi)^2 = 0.$$

Из последнего уравнения найдем выражение для функции f

$$f_\xi(x) = \exp(\lambda_1 + \lambda_2 x + \lambda_3 (x - m_\xi)^2 - 1).$$

Из дополнительных условий получим

$$\begin{aligned} \int_{-\infty}^{\infty} \exp(\lambda_1 + \lambda_2 x + \lambda_3 (x - m_\xi)^2 - 1) dx &= 1, \\ \int_{-\infty}^{\infty} x \exp(\lambda_1 + \lambda_2 x + \lambda_3 (x - m_\xi)^2 - 1) dx &= m_\xi, \\ \int_{-\infty}^{\infty} (x - m_\xi)^2 \exp(\lambda_1 + \lambda_2 x + \lambda_3 (x - m_\xi)^2 - 1) dx &= \sigma_\xi^2. \end{aligned}$$

Легко показать, что

$$\lambda_1 = 1 - \ln(\sqrt{2\pi}\sigma_\xi), \quad \lambda_2 = 0, \quad \lambda_3 = -(2\sigma_\xi^2)^{-1}.$$

Получаем

$$\hat{f}_\xi(x) = \frac{1}{\sigma_\xi \sqrt{2\pi}} \exp\left\{-\frac{(x - m_\xi)^2}{2\sigma_\xi^2}\right\}$$

нормальное распределение, которое доставляет максимум энтропии при фиксированных двух первых моментах m_ξ и σ_ξ^2 .

Полная взаимная информация, содержащаяся в непрерывных системах ξ and η определяется следующим образом

$$I_{\eta \leftrightarrow \xi} = \iint_{-\infty}^{\infty} f_{\xi\eta}(x, y) \log \frac{f_{\xi\eta}(x, y)}{f_{\xi}(x)f_{\eta}(y)} dx dy$$

или

$$I_{\eta \leftrightarrow \xi} = M[\log(f_{\xi\eta}/f_{\xi}f_{\eta})].$$

Если случайные величины ξ и η независимы, тогда полная взаимная информация равна нулю:

$$I_{\eta \leftrightarrow \xi} = 0.$$

Пример: Рассмотрим простую модель

$$\eta = \xi + \zeta,$$

где η – полученный (измеренный) сигнал, ξ – переданный сигнал, с нормальной плотностью $N(m_{\xi} = 0, \sigma_{\xi}^2)$, ζ – аддитивный шум с нормальной плотностью $(m_{\zeta}, \sigma_{\zeta}^2)$, ξ и ζ – независимы. Требуется найти количество информации, содержащееся в измеренном сигнале η относительно переданного сигнала ξ . По определению:

$$I_{\eta \leftrightarrow \xi} = M \left[\log \frac{f_{\xi\eta}}{f_{\xi}f_{\eta}} \right] = M \left[\log \frac{f_{\eta/\xi}}{f_{\eta}} \right] = M \varkappa,$$

где $\varkappa = \log(f_{\eta/\xi}/f_{\eta})$. Можно записать $\sigma_{\eta}^2 = \sigma_{\xi}^2 + \sigma_{\zeta}^2$ and

$$f_{\eta}(y) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{\xi}^2 + \sigma_{\zeta}^2}} \exp \left\{ -\frac{1}{2} \frac{y^2}{(\sigma_{\xi}^2 + \sigma_{\zeta}^2)} \right\},$$

$$f_{\eta/\xi}(y/x) = \frac{1}{\sqrt{2\pi}\sigma_{\eta}\sqrt{1-r^2}} \exp \left\{ -\frac{1}{2(1-r^2)} \left[\frac{y}{\sigma_{\eta}} - r \frac{x}{\sigma_{\xi}} \right]^2 \right\}.$$

Найдем коэффициент корреляции r

$$r = \frac{1}{\sigma_{\eta}\sigma_{\xi}} M(\eta \cdot \xi) = \frac{1}{\sigma_{\eta}\sigma_{\xi}} [M(\xi^2) + M(\xi\zeta)] = \frac{\sigma_{\xi}^2}{\sigma_{\eta}^2}.$$

Подставляя r в $f_{\eta/\xi}(y/x)$ получаем

$$f_{\eta/\xi}(y/x) = \frac{1}{\sqrt{2\pi}\sigma_{\zeta}} \exp \left\{ -\frac{(y-x)^2}{2\sigma_{\zeta}^2} \right\} = \frac{1}{\sqrt{2\pi}\sigma_{\zeta}} \exp \left\{ -\frac{z^2}{2\sigma_{\zeta}^2} \right\}.$$

Получим выражение для случайной величины \varkappa

$$\varkappa = \log \frac{\sqrt{\sigma_{\xi}^2 + \sigma_{\zeta}^2}}{\sigma_{\zeta}} + \frac{1}{\ln 2} \left[\frac{\zeta^2}{2\sigma_{\zeta}^2} - \frac{\eta^2}{2(\sigma_{\xi}^2 + \sigma_{\zeta}^2)} \right].$$

После вычисления математического ожидания случайной величины \varkappa , получаем

$$I_{\eta \leftrightarrow \xi} = M\varkappa = \log \frac{\sqrt{\sigma_\xi^2 + \sigma_\zeta^2}}{\sigma_\zeta} + \frac{1}{\ln 2} \left[\frac{M\zeta^2}{2\sigma_\zeta^2} - \frac{M\eta^2}{2(\sigma_\xi^2 + \sigma_\zeta^2)} \right].$$

С учетом $M\zeta^2 = \sigma_\zeta^2$, $M\eta^2 = \sigma_\xi^2 + \sigma_\zeta^2$ записываем:

$$I_{\eta \leftrightarrow \xi} = \log \frac{\sqrt{\sigma_\xi^2 + \sigma_\zeta^2}}{\sigma_\zeta}.$$

2.5 Информация Фишера

Пусть случайная величина ξ имеет плотность вероятностей $f_\xi(x, \theta)$, где θ – параметр, и $f_\xi(x, \theta)$ дифференцируема по θ . *Информация Фишера* о неизвестном параметре θ , содержащаяся в случайной величине ξ определяется следующим образом

$$I_\xi^{(F)}(\theta) = M \left[\left(\frac{\partial \log f_\xi(x, \theta)}{\partial \theta} \right)^2 \right].$$

Для случайного вектора $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$, и параметрического вектора $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_s)$ определяют *матрицу Фишера* с элементами

$$I_{ss'}^{(F)}(\vec{\theta}) = M \left[\frac{\partial \log f_{\vec{\xi}}(x_1, x_2, \dots, x_n, \vec{\theta})}{\partial \theta_s} \frac{\partial \log f_{\vec{\xi}}(x_1, x_2, \dots, x_n, \vec{\theta})}{\partial \theta_{s'}} \right].$$

Свойства информации Фишера.

1. Пусть случайные величины ξ_1 и ξ_2 независимы, при этом информация Фишера относительно параметра θ – соответственно $I_{\xi_1}^{(F)}(\theta)$ и $I_{\xi_2}^{(F)}(\theta)$. Информация Фишера о параметре θ , содержащаяся в паре ξ_1, ξ_2 равна

$$I_{\xi_1 \xi_2}^{(F)}(\theta) = I_{\xi_1}^{(F)}(\theta) + I_{\xi_2}^{(F)}(\theta).$$

Доказательство:

$$\begin{aligned} I_{\xi_1 \xi_2}^{(F)}(\theta) &= M \{ \partial [\log f_{\xi_1}(x, \theta) \cdot f_{\xi_2}(x, \theta)] / \partial \theta \}^2 = \\ &= M [(\partial \log f_{\xi_1}(x, \theta) / \partial \theta)^2] + M [(\partial \log f_{\xi_2}(x, \theta) / \partial \theta)^2] + \\ &+ 2M [\partial \log f_{\xi_1}(x, \theta) / \partial \theta \cdot \partial \log f_{\xi_2}(x, \theta) / \partial \theta] = \\ &= I_{\xi_1}^{(F)}(\theta) + I_{\xi_2}^{(F)}(\theta). \end{aligned}$$

Учитываем, что

$$\begin{aligned} M [\partial \log f_{\xi_1}(x, \theta) / \partial \theta \cdot \partial \log f_{\xi_2}(x, \theta) / \partial \theta] &=, \\ = M [\partial \log f_{\xi_1}(x, \theta) / \partial \theta] \cdot M [\partial \log f_{\xi_2}(x, \theta) / \partial \theta] &= 0 \end{aligned}$$

т.к.

$$\int_{-\infty}^{\infty} f_{\xi}(x, \theta) \frac{1}{f_{\xi}(x, \theta)} \frac{\partial f_{\xi}(x, \theta)}{\partial \theta} dx = \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} f_{\xi}(x, \theta) dx = 0.$$

2. Пусть случайные величины $\xi_1, \xi_2, \dots, \xi_n$ независимы. Информация Фишера, содержащаяся в каждой случайной величине – $I_{\xi}^{(F)}(\theta)$. Тогда информация Фишера равна $I_{\vec{\xi}}^{(F)}(\theta)$ equals

$$I_{\vec{\xi}}^{(F)} = nI_{\xi}^{(F)}(\theta)$$

- а) Для случайных величин ξ_1 и ξ_2 и векторного параметра $\vec{\theta}$ информационная матрица Фишера удовлетворяет соотношению

$$I_{\xi_1 \xi_2}^{(F)}(\vec{\theta}) = I_{\xi_1}^{(F)}(\vec{\theta}) + I_{\xi_2}^{(F)}(\vec{\theta}).$$

- б) В случае случайного вектора $\vec{\xi} = [\xi_1, \xi_2, \dots, \xi_n]$ и вектора параметров $\vec{\theta}$ справедливо

$$I_{\vec{\xi}}^{(F)}(\vec{\theta}) = nI_{\xi}^{(F)}(\vec{\theta}).$$

3. Пусть $T = T(\xi)$ – функция от случайной переменной ξ , тогда удовлетворяется следующее неравенство

$$I_{\xi}^{(F)}(\vec{\theta}) - I_{T(\xi)}^{(F)}(\vec{\theta}) \geq 0.$$

В случае одного параметра

$$I_{\xi}^{(F)} \geq I_{T(\xi)}^{(F)},$$

количество информации Фишера, содержащееся в случайной величине ξ всегда больше или равно количеству информации Фишера содержащегося в функции $T(\xi)$.

4. Можно представить информацию Фишера в виде:

$$I_{\xi}^{(F)}(\theta) = M[-\partial^2 \log f_{\xi}(x, \theta) / \partial \theta^2],$$

поскольку легко показать, что

$$\frac{\partial^2 \log f}{\partial \theta^2} = -\frac{1}{f^2} \left(\frac{\partial f}{\partial \theta} \right)^2 + \frac{1}{f} \frac{\partial^2 f}{\partial \theta^2}, \quad \left(\frac{1}{f} \frac{\partial f}{\partial \theta} \right)^2 = \left(\frac{\partial \log f}{\partial \theta} \right)^2$$

и

$$\begin{aligned} M\left(\frac{1}{f} \frac{\partial^2 f}{\partial \theta^2}\right) &= \int_{-\infty}^{\infty} f_{\xi}(x, \theta) \frac{1}{f_{\xi}(x, \theta)} \frac{\partial^2 f_{\xi}(x, \theta)}{\partial \theta^2} dx = \\ &= \frac{\partial^2}{\partial \theta^2} \int_{-\infty}^{\infty} f_{\xi}(x, \theta) dx = 0. \end{aligned}$$

В случае вектора параметров $\vec{\theta}$ получаем

$$I_{ss'}^{(F)}(\vec{\theta}) = M[-\partial^2 \log f(x, \vec{\theta}) / \partial \theta_s \partial \theta_{s'}].$$

Пример: Пусть случайная величина ξ имеет нормальное распределение $\xi \in N(m_\xi, \sigma_\xi^2)$. Пусть математическое ожидание m_ξ неизвестно, а дисперсия σ_ξ^2 известна. Информация Фишера относительно математического ожидания равна

$$\begin{aligned}
 I_\xi^{(F)}(m_\xi) &= - \int_{-\infty}^{\infty} \frac{\partial^2}{\partial m_\xi^2} \left[-\frac{(x - m_\xi)^2}{2\sigma_\xi^2} - \ln \sigma_\xi \sqrt{2\pi} \right] \times \\
 &\quad \times \frac{1}{\sqrt{2\pi}\sigma_\xi} \exp\left[-\frac{(x - m_\xi)^2}{2\sigma_\xi^2}\right] dx = \\
 &= \int_{-\infty}^{\infty} \frac{1}{\sigma_\xi^2 \sqrt{2\pi}\sigma_\xi} \exp\left[-\frac{(x - m_\xi)^2}{2\sigma_\xi^2}\right] dx = \frac{1}{\sigma_\xi^2}.
 \end{aligned}$$

Глава 3

Свойства случайных функций

Определение: Пусть t – неслучайная величина. Случайная функция $\xi(t)$ называется *случайным процессом* если t – непрерывная величина и $\xi(t_i)$ называется *временным рядом*, если t дискретна.

Предположим, что имеется n независимых экспериментов. В каждом эксперименте получаем n реализаций случайной функции $\xi(t)$ (см. рис. 3.1):

$$x_1(t), x_2(t), \dots, x_n(t).$$

Случайная функция $\xi(t)$ в момент $t = t_1$ является случайной величиной $\xi_1 =$

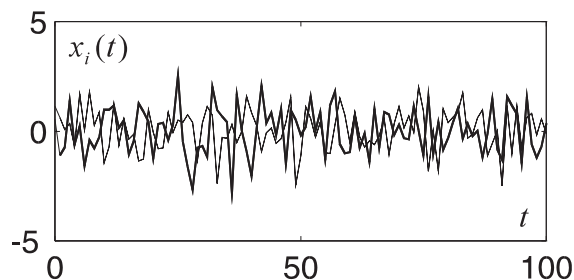


Рис. 3.1: Три реализации случайного процесса

$\xi(t_1)$ и описывается плотностью вероятностей $f(x_1/t_1)$.

В двух точках t_1 и t_2 случайные величины $\xi_1 = \xi(t_1)$ и $\xi_2 = \xi(t_2)$ описываются плотностью $f(x_1, x_2/t_1, t_2)$.

Случайная функция $\xi(t)$ полностью определяется плотностью вероятностей

$$f(x_1, x_2, \dots, x_n/t_1, t_2, \dots, t_n).$$

Два первых момента играют очень важную роль в приложениях.

3.1 Математическое ожидание и дисперсия

Математическое ожидание определяется следующим образом

$$m_{\xi_1} = M[\xi(t_1)] = \langle \xi(t_1) \rangle$$

или, для произвольного t ,

$$m_\xi(t) = \int_{-\infty}^{\infty} x f(x/t) dx.$$

Дисперсия определяется как

$$D[\xi(t)] = M[(\xi(t) - m_\xi(t))^2] = \langle (\xi(t) - m_\xi(t))^2 \rangle.$$

Корреляционная функция (автокорреляционная функция) определяется как

$$\begin{aligned} R(t_1, t_2) &= \text{cov}(\xi(t_1), \xi(t_2)) = M[(\xi(t_1) - m_\xi(t_1))(\xi(t_2) - m_\xi(t_2))] = \\ &= \langle (\xi(t_1) - m_\xi(t_1))(\xi(t_2) - m_\xi(t_2)) \rangle \end{aligned}$$

или, используя оператор математического ожидания:

$$\begin{aligned} D[\xi(t)] &= \int_{-\infty}^{\infty} (x(t) - m_\xi(t))^2 f(x/t) dx, \\ R(t_1, t_2) &= \iint_{-\infty}^{\infty} (x(t_1) - m_\xi(t_1))(x(t_2) - m_\xi(t_2)) f(x_1, x_2/t_1, t_2) dx_1 dx_2. \end{aligned}$$

Среднее по времени (по параметру t) случайной функции $\xi(t)$ определяется:

$$\langle \xi(t) \rangle_T = \lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T \xi(t) dt.$$

3.2 Свойства случайных функций

3.2.1 Стационарность

Дадим два определения стационарности.

А. Стационарность в узком смысле. Говорят, что случайная функция $\xi(t)$ является *стационарной (в узком смысле)*, если ее плотность вероятности не зависит от сдвига по времени:

$$f(x_1, \dots, x_n/t_1, \dots, t_n) = f(x_1, \dots, x_n/t_1 + t_0, \dots, t_n + t_0)$$

для любого t_0 .

В частном случае $n = 1$ и $n = 2$ имеем соответственно:

$$f(x_1/t_1) = f(x_1/0) = f(x_1)$$

и

$$f(x_1, x_2/t_1, t_2) = f(x_1, x_2/0, \tau) \quad \tau = t_2 - t_1,$$

т.е. одномерная плотность не зависит от времени и двумерная плотность зависит только от разности времен.

В. Стационарность в широком смысле. Говорят, что случайная функция $\xi(t)$ является *стационарной в широком смысле*, если ее математическое ожидание и дисперсия постоянны и ее автокорреляционная функция зависит только от разности $\tau = t_2 - t_1$

$$m_\xi(t) = \int_{-\infty}^{\infty} x f(x) dx = const,$$

$$D[\xi(t)] = \int_{-\infty}^{\infty} (x(t) - m_\xi(t))^2 f(x) dx = const,$$

$$R(t_1, t_2) = \int_{-\infty}^{\infty} (x(t_1) - m_\xi(t_1))(x(t_2) - m_\xi(t_2)) f(x_1, x_2/t_1, t_2) dx_1 dx_2 = R(t_2 - t_1) = R(\tau),$$

$$m_\xi = const, \quad D[\xi(t)] = const, \quad R(t_1, t_2) = R(t_2 - t_1) = R(\tau).$$

Примеры нестационарных случайных процессов представлены на рис. 3.2.

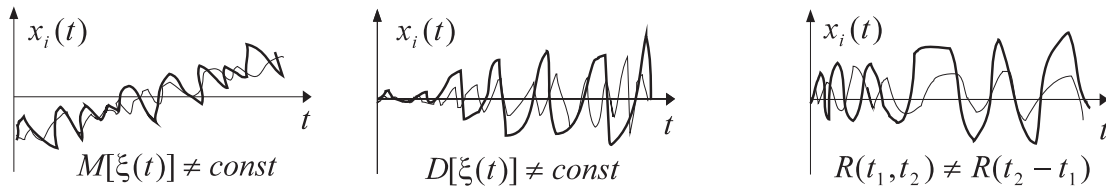


Рис. 3.2: Примеры нестационарных процессов.

3.2.2 Эргодичность

Понятие *эргодичности* связано с определением статистик процесса $\xi(t)$ по его реализации. from a signal observation. Случайный процесс $\xi(t)$ называют эргодическим в наиболее общем смысле, если (с вероятностью 1) все его статистики могут быть определены по одной его реализации. Статистический процесс $\xi(t)$ является эргодическим, если усреднение по времени равно усреднению по ансамблям

$$P[\langle \xi(t) \rangle_T = M[\xi(t)]] = 1.$$

(т.е. можно вычислить $M[\xi(t)]$ в фиксированной точке t_i с использованием множества реализаций случайного процесса (см. рис. 3.3а) или вычислить эту величину усреднением во временной области с использованием только одной реализации (см. рис. 3.3б) случайного процесса.)

3.3 Свойства автокорреляционных функций

1. Автокорреляционная функция – нечетная функция

$$R(\tau) = R(-\tau).$$

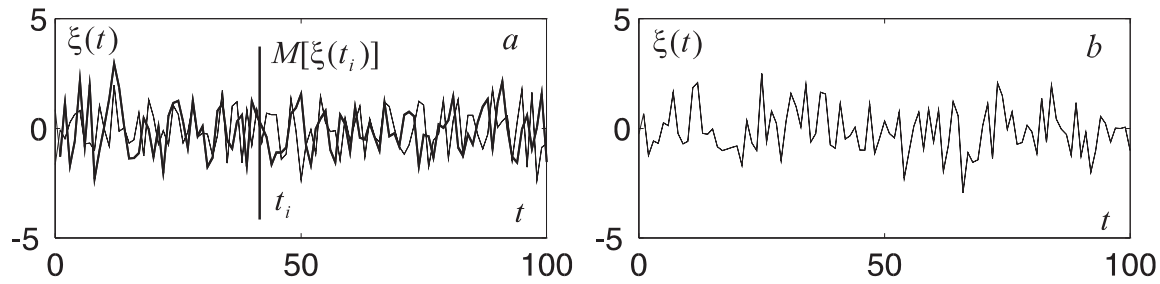


Рис. 3.3: Иллюстрация эргодичности.

Это свойство следует непосредственно из определения.

2. Величина автокорреляционной функции меньше или равна дисперсии:

$$D[\xi(t)] = R(0) \geq R(\tau).$$

3. Отправляясь от произведения величины $[\xi(t) - m_\xi(t)]$ на произвольную неслучайную функцию $z(t)$ можно получить следующее неравенство

$$\int_a^b \int_a^b z(t_1)z(t_2)R(t_2 - t_1)dt_1dt_2 \geq 0.$$

Примеры аппроксимации автокорреляционных функций представлены на рис. 3.4.

4. Пусть $\varphi(t)$ – неслучайная функция, а $\xi(t)$ – случайная. Рассмотрим функцию $\eta(t) = \xi(t) + \varphi(t)$. Учтем, что математическое ожидание η можно записать в виде

$$m_\eta(t) = m_\xi(t) + \varphi(t),$$

получим автокорреляционную функцию $R_\eta(t_1, t_2)$

$$\begin{aligned} R_\eta(t_1, t_2) &= M[(\eta(t_1) - m_\eta(t_1))(\eta(t_2) - m_\eta(t_2))] = \\ &= M[(\xi(t_1) - m_\xi(t_1))(\xi(t_2) - m_\xi(t_2))] = R_\xi(t_1, t_2). \end{aligned}$$

3.4 Действие линейного оператора на случайную функцию

Рассмотрим линейную систему с импульсным (вещественным) откликом $h(t)$ и входным сигналом $\xi(t)$. Выходной сигнал $\eta(t)$ (см. рис. 3.5) дается формулой

$$\eta(t) = H[\xi(t)] = \int_{-\infty}^{\infty} \xi(t - \tau)h(\tau)d\tau.$$

Свойства линейного оператора H .

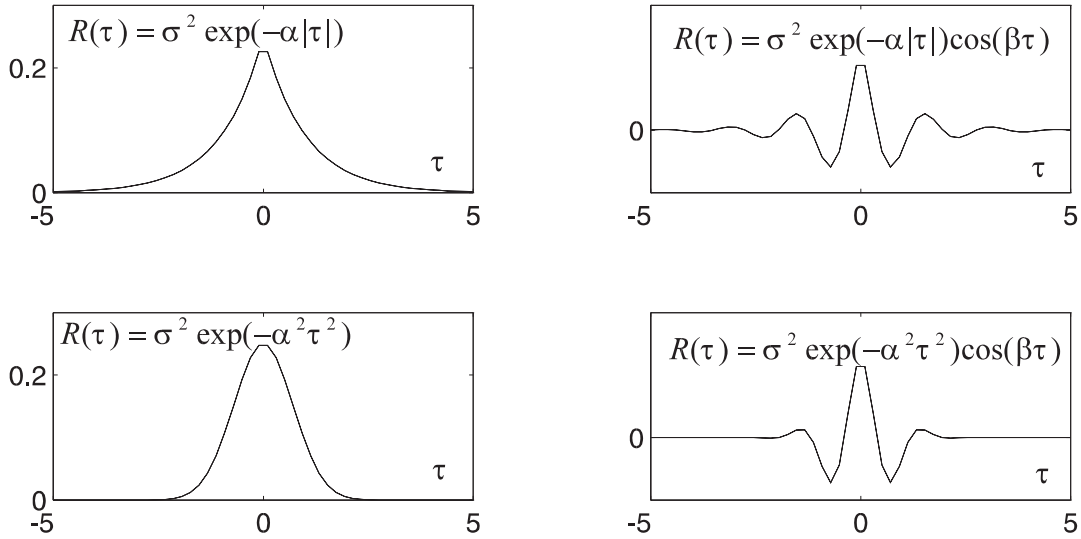


Рис. 3.4: Примеры аппроксимации автокорреляционных функций ($\sigma = 0.5$, $\alpha = 1.0$, $\beta = 4.0$).

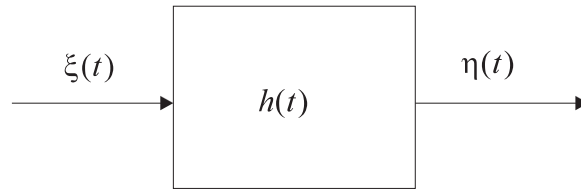


Рис. 3.5: Схема действия оператора H .

1. $H[c\xi(t)] = cH[\xi(t)]$.
2. $H[\xi_1(t) + \xi_2(t)] = H\xi_1(t) + H\xi_2(t)$.
3. $\eta(t) = H[\xi(t)]$, $m_\eta(t) = M[H[\xi(t)]] = H[M[\xi(t)]] = Hm_\xi(t)$,
 $R_\eta(t_1, t_2) = M[(\eta(t_1) - m_\eta(t_1))(\eta(t_2) - m_\eta(t_2))] =$
 $= M\{(H[\xi(t_1)] - H[m_\xi(t_1)])(H[\xi(t_2)] - H[m_\xi(t_2)])\} =$
 $= H_{t_1}H_{t_2}M[(\xi(t_1) - m_\xi(t_1))(\xi(t_2) - m_\xi(t_2))] = H_{t_1}H_{t_2}R_\xi(t_1, t_2)$.

3.5 Взаимокорреляционная функция

Рассмотрим двумерный процесс $\xi(t)$ и $\eta(t)$.

Взаимокорреляционная функция определяется следующим образом

$$R_{\xi\eta}(t_1, t_2) = M[(\xi(t_1) - m_\xi(t_1))(\eta(t_2) - m_\eta(t_2))].$$

В случае вещественных случайных функций получаем

$$R_{\xi\eta}(t_1, t_2) = R_{\eta\xi}(t_2, t_1)$$

и для стационарных $\xi(t)$ и $\eta(t)$:

$$R_{\xi\eta}(\tau) = R_{\eta\xi}(-\tau).$$

В согласии с определением взаимокорреляционной функции можно записать

$$R_{\xi\eta}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x(t_1) - m_{\xi}(t_1))(y(t_2) - m_{\eta}(t_2)) f_{\xi\eta}(x, y/t_1, t_2) dx dy$$

и, используя неравенство Коши–Буняковского, получаем

$$\begin{aligned} & \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x(t_1) - m_{\xi}(t_1))(y(t_2) - m_{\eta}(t_2)) f(x, y/t_1, t_2) dx dy \right]^2 \leq \\ & \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x(t_1) - m_{\xi}(t_1))^2 f(x, y/t_1, t_2) dx dy \times \\ & \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y(t_2) - m_{\eta}(t_2))^2 f(x, y/t_1, t_2) dx dy \Rightarrow \\ & [R_{\xi\eta}(t_1, t_2)]^2 \leq \sqrt{R_{\xi}(t_1, t_1) R_{\eta}(t_2, t_2)}. \end{aligned}$$

После нормировки будем иметь нормированную взаимокорреляционную функцию

$$r_{\xi\eta}(t_1, t_2) = \frac{R_{\xi\eta}(t_1, t_2)}{\sqrt{R_{\xi}(t_1, t_1)} \sqrt{R_{\eta}(t_2, t_2)}}, \quad -1 \leq r_{\xi\eta}(t_1, t_2) \leq 1$$

и

$$R_{\xi\eta}(\tau) \leq \sqrt{D_{\xi} D_{\eta}}.$$

3.6 Теорема Винера–Хинчина и спектр мощности

1. Пусть существуют интегралы

$$\langle |\xi| \rangle_T = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |x(t)| dt$$

и

$$\langle |\xi|^2 \rangle_T = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |x(t)|^2 dt.$$

2. Пусть $\xi(t)$ – случайная функция с ограниченной вариацией. Корреляционная функция

$$R_\xi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \xi(t)\xi(t+\tau)dt$$

существует всегда, если существует $\langle |\xi|^2 \rangle_T$ exists. Взаимокорреляционная функция

$$R_{\xi\eta}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \xi(t)\eta(t+\tau)dt$$

существует всегда, если существуют $\langle |\xi|^2 \rangle_T$ и $\langle |\eta|^2 \rangle_T$.

Спектральная плотность $R_\xi(\omega)$ для функции $\xi(t)$ и взаимная спектральная плотность $R_{\xi\eta}(\omega)$ для функций $\xi(t)$ и $\eta(t)$ определяются соотношениями Винера-Хинчина

$$\begin{aligned} R_\xi(\omega) &= \int_{-\infty}^{\infty} R_\xi(\tau) \exp\{-i\omega\tau\}d\tau = 2 \int_0^{\infty} R_\xi(\tau) \cos \omega\tau d\tau, \\ R_\xi(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} R_\xi(\omega) \exp\{i\omega\tau\}d\omega = \frac{1}{\pi} \int_0^{\infty} R_\xi(\omega) \cos \omega\tau d\omega, \\ R_{\xi\eta}(\omega) &= \int_{-\infty}^{\infty} R_{\xi\eta}(\tau) \exp\{-i\omega\tau\}d\tau, \\ R_{\xi\eta}(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{\xi\eta}(\omega) \exp\{i\omega\tau\}d\omega. \end{aligned}$$

Для вещественной случайной функции $\xi(t)$ спектральная плотность является нечетной

$$R_\xi(\omega) = R_\xi(-\omega).$$

3.7 Пример корреляционных функций и спектров мощности

Пример 1: Графическое представление корреляционной функции

$$R_\xi(\tau) = \sigma_\xi^2 \exp\{-\alpha|\tau|\} \quad (3.1)$$

и ее спектра

$$R_\xi(\omega) = \sigma_\xi^2 \int_{-\infty}^{\infty} \exp\{-i\omega\tau - \alpha|\tau|\}d\tau = \sigma_\xi^2 \left[\int_0^{\infty} \exp\{-(i\omega + \alpha)\tau\}d\tau + \right.$$

$$+ \int_{-\infty}^0 \exp\{(\alpha - i\omega)\tau\} d\tau] = \sigma_\xi^2 \left[\frac{1}{\alpha + i\omega} + \frac{1}{\alpha - i\omega} \right] = 2\sigma_\xi^2 \frac{\alpha}{\alpha^2 + \omega^2} \quad (3.2)$$

– см. рис. 3.6а и 3.6b соответственно.

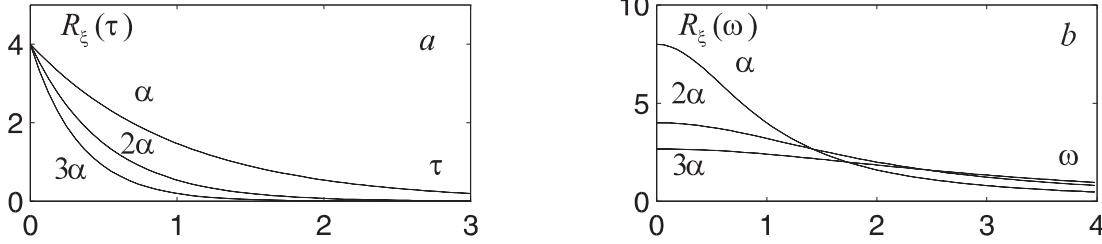


Рис. 3.6: Графическое представление корреляционной функции (3.1) (а) и ее спектра (3.2)(b) ($\sigma = 2$, $\alpha = 1$).

Пример 2: $R_\xi(\tau) = \sigma_\xi^2 \exp\{-\alpha|\tau|\} \cos \omega_0 \tau$.

Рассмотрим следующее представление $\cos \omega_0 \tau$

$$\cos \omega_0 \tau = (\exp\{i\omega_0 \tau\} + \exp\{-i\omega_0 \tau\})/2.$$

Выполним преобразование Фурье и получим:.

$$\begin{aligned} R_\xi(\omega) &= \frac{\sigma_\xi^2}{2} \left[\int_{-\infty}^{\infty} \exp\{-i\omega\tau - \alpha|\tau| + i\omega_0\tau\} d\tau + \right. \\ &\quad \left. + \int_{-\infty}^{\infty} \exp\{-i\omega\tau - \alpha|\tau| - i\omega_0\tau\} d\tau \right] = \\ &= \sigma_\xi^2 \left[\frac{\alpha}{(\omega - \omega_0)^2 + \alpha^2} + \frac{\alpha}{(\omega + \omega_0)^2 + \alpha^2} \right]. \end{aligned}$$

Example 3: $R_\xi(\tau) = \sigma_\xi^2 \exp\{-\alpha^2 \tau^2\} \cos \omega_0 \tau$.

Рассмотрим представление $\cos \omega_0 \tau$

$$\cos \omega_0 \tau = (\exp\{i\omega_0 \tau\} + \exp\{-i\omega_0 \tau\})/2$$

после выполнения преобразования Фурье получаем

$$\begin{aligned} R_\xi(\omega) &= \frac{\sigma_\xi^2}{2} \left[\int_{-\infty}^{\infty} \exp\{-\alpha^2 \tau^2 + i\omega_0 \tau - i\omega \tau\} d\tau + \right. \\ &\quad \left. + \int_{-\infty}^{\infty} \exp\{-\alpha^2 \tau^2 - i\omega_0 \tau - i\omega \tau\} d\tau \right]. \end{aligned}$$

Каждый интеграл можно свести к интегралу Пуассона, представляя экспоненту в виде

$$\left[\left(\alpha\tau + i\frac{\omega}{2\alpha} \mp i\frac{\omega_0}{2\alpha} \right)^2 + \frac{1}{4\alpha^2}(\omega \pm \omega_0)^2 \right].$$

Вычисли подобные интегралы получим

$$R_\xi(\omega) = \frac{\sigma_\xi^2 \sqrt{\pi}}{2\alpha} \left[\exp\left\{ \frac{-(\omega + \omega_0)^2}{4\alpha^2} \right\} + \exp\left\{ \frac{-(\omega - \omega_0)^2}{4\alpha^2} \right\} \right].$$

3.8 Оценивание численных характеристик случайных функций

Пусть реализация случайной функции зарегистрирована с шагом Δt :

$$x_i(t_j), \quad j = 1, \dots, m; \quad i = 1, \dots, n.$$

Оценка математического ожидания:

$$\hat{m}_\xi(t_j) = \frac{1}{n} \sum_{i=1}^n x_i(t_j).$$

Оценка корреляционной функции может быть представлена как:

$$\hat{R}_\xi(t_j, t_k) = \frac{1}{n-1} \sum_{i=1}^n [(x_i(t_j) - \hat{m}_\xi(t_j))(x_i(t_k) - \hat{m}_\xi(t_k))]$$

или

$$\hat{R}_\xi(t_j, t_k) = \frac{n}{n-1} \sum_{i=1}^n x_i(t_j)x_i(t_k) - \frac{n}{n-1} m_\xi(t_j)m_\xi(t_k).$$

В случае стационарного эргодического процесса оценку математического ожидания можно искать как

$$\hat{m}_\xi = (1/m) \sum_{j=1}^m x(t_j)$$

(для дискретного случая)

$$\hat{m}_\xi = \frac{1}{T} \int_0^T x(t) dt,$$

(для непрерывного случая, при этом оценка корреляционной функции записывается следующим образом

$$\hat{R}_\xi(\tau) = \frac{1}{m-l-1} \sum_{j=1}^{m-l} (x(t_j) - \hat{m}_\xi)(x(t_j + \tau) - \hat{m}_\xi)$$

(для дискретного случая),

$$\hat{R}_\xi(\tau) = \frac{1}{T - \tau} \int_0^{T-\tau} [x(t) - \hat{m}_\xi][x(t + \tau) - \hat{m}_\xi] dt$$

(для непрерывного случая).

Глава 4

Элементы математической статистики

4.1 Оценивание

Пусть $\vec{\xi} = (x_1, x_2, \dots, x_n)$ – случайный вектор полученный регистрацией исходов случайного эксперимента. Величины, формирующие такой вектор или их реализации называются *наблюдениями* или *измерениями*.

Пусть функция $g(x_1, \dots, x_n)$ – оценка.

Пример 1: Оценка среднего значения

$$\bar{x} = (1/n) \sum_{i=1}^n x_i.$$

Пример 2: Оценка дисперсии

$$s^2 = (1/n) \sum_{i=1}^n (x_i - \bar{x})^2.$$

Пример 3: Оценка коэффициента вариации

$$v = \frac{s}{\bar{x}}.$$

Пусть вектор $\vec{\xi}$ принадлежит плотности вероятностей

$$\vec{\xi} \in f(\vec{x}, \vec{\theta}), \quad \theta \in \Omega,$$

где $\vec{\theta}$ – неизвестный векторный параметр. Основные моменты связанные с решением задач математической статистики:

1. Класс распределений

$$\mathcal{F} = \{f(\vec{x}, \vec{\theta}), \vec{\theta} \in \Omega\}$$

для случайного вектора $\vec{\xi}$.

2. Структура пространства D решений $d, d \in D$.
3. Пространства наблюдений (данные) $\vec{x} \in X$.
4. Пространство неизвестных (искомых) параметров $\vec{\theta} \in \Omega$

Связи между X , D и Ω представлены на рис. 4.1. Оценивание неизвестного век-

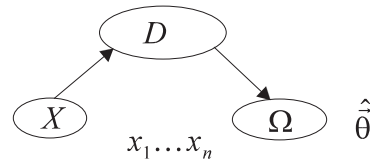


Рис. 4.1: Связи между пространствами решений, наблюдений и параметров.

тора $\vec{\theta}$ с использованием функции $g(x_1, \dots, x_n)$, зависящей от наблюдений \vec{x} называется *точечным оцениванием*. Функция $g(\vec{x})$ называется *оценка*

$$\hat{\theta}_n = g(x_1, \dots, x_n).$$

4.2 Состоятельные оценки

Оценка называется *состоятельной*, если при стремлении выборки наблюдений к бесконечности, оценка стремится к истинному значению по вероятности

$$P(|\hat{\theta}_n - \theta_0| \leq \varepsilon) \geq 1 - \eta,$$

где $\varepsilon > 0$, $\eta > 0$ (см. рис. 4.2).

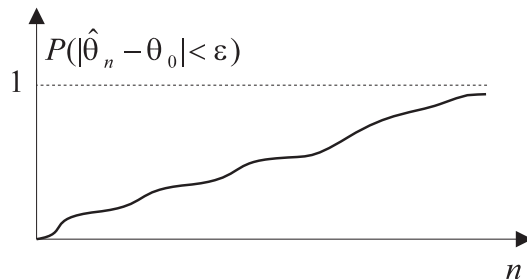


Рис. 4.2: Иллюстрация сходимости по вероятности.

4.3 Несмещенное оценивание

Пусть $\hat{\theta}_n$ – оценка параметра (с истинным значением) θ_0 . Математическое ожидание $M[\hat{\theta}_n - \theta_0]$ ошибки оценки называется *смещением*. Если смещение равно нулю, то такая оценка называется *несмещенной*. При этом выполняется

$$b_n(\hat{\theta}) = M(\hat{\theta}_n) - \theta_0 = 0, \quad M[\hat{\theta}] = \theta_0.$$

Асимптотически несмещенная оценка:

$$M(\hat{\theta}_n) \xrightarrow{n \rightarrow \infty} \theta_0$$

(иллюстрации – см. рис. 4.3).

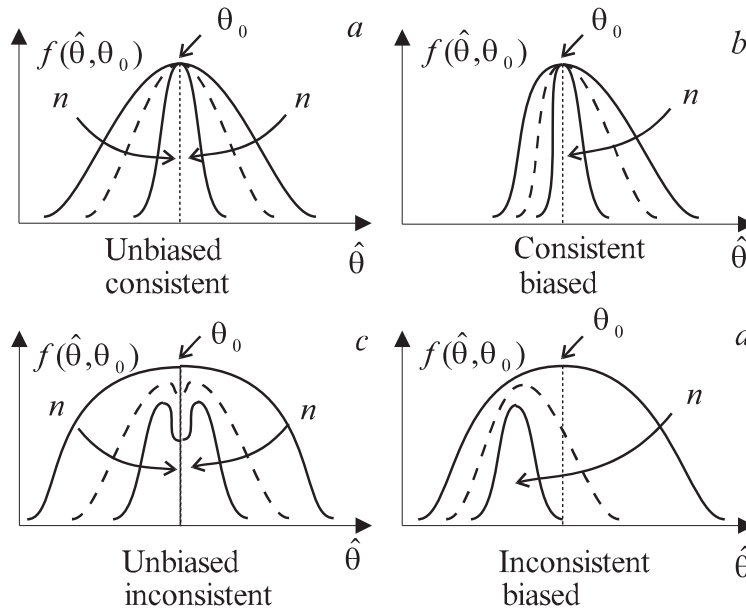


Рис. 4.3: Иллюстрации различных типов оценок.

Пример: Пусть (x_1, x_2, \dots, x_n) – выборка нормально распределенной случайной величины $x_i \in N(m_\xi, \sigma_\xi^2)$. Найдем оценку \hat{S}^2 дисперсии σ^2 .

Вычислим выборочную дисперсию

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2.$$

Математическое ожидание:

$$M(s^2) = \frac{1}{n} \sum_{i=1}^n M(x_i^2) - M(\bar{x}^2) = \left(1 - \frac{1}{n}\right) \sigma^2.$$

Следовательно, несмещенная оценка дисперсии имеет вид

$$\hat{S}^2 = \frac{n}{n-1} s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad M(\hat{S}^2) = \frac{n}{n-1} M(s^2) = \sigma^2.$$

4.4 Неравенство Рао–Крамера. Эффективность

Нижняя оценка дисперсии оцениваемой величины определяется неравенством Рао–Крамера.

Пусть x_1, x_2, \dots, x_n – выборка, отвечающая случайной величине с плотностью вероятностей

$$L(\vec{x}, \theta) = \prod_{i=1}^n f(x_i, \theta).$$

Пусть $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ – оценка. Плотность вероятности $L(\vec{x}, \theta)$ нормирована на единицу

$$\int \dots \int L(x_1, x_2, \dots, x_n, \theta) dx_1, \dots, dx_n = 1.$$

Выполнив дифференцирование по параметру θ получим

$$\int \dots \int \frac{\partial L}{\partial \theta} dx_1, \dots, dx_n = 0. \quad (4.1)$$

В силу определения несмещенности оценки $\hat{\theta}$

$$M\hat{\theta} = \theta$$

или

$$\int \dots \int \hat{\theta} L dx_1, \dots, dx_n = \theta, \quad \int \dots \int \hat{\theta}(x_1, \dots, x_n) \frac{\partial L}{\partial \theta} dx_1, \dots, dx_n = 1. \quad (4.2)$$

С учетом (4.1) и (4.2) можно записать

$$\int \dots \int [\hat{\theta} - \theta] \frac{\partial L(x_1, \dots, x_n, \theta)}{\partial \theta} dx_1, \dots, dx_n = 1$$

или

$$\int \dots \int [\hat{\theta} - \theta] \left[\frac{1}{L} \frac{\partial L}{\partial \theta} \right] L dx_1, \dots, dx_n = 1.$$

Последнее соотношение можно представить в виде

$$M \left[(\hat{\theta} - \theta) \left(\frac{1}{L} \frac{\partial L}{\partial \theta} \right) \right] = 1.$$

Вводя обозначения

$$\hat{\theta} - \theta = V \quad \text{и} \quad \frac{1}{L} \frac{\partial L}{\partial \theta} = W$$

и, используя аналог неравенства Коши–Буняковского, получаем

$$M(V^2) \cdot M(W^2) \geq [M(V \cdot W)]^2$$

или

$$M(V^2) \cdot M(W^2) \geq 1, \quad M(\hat{\theta} - \theta)^2 \cdot M \left[\frac{1}{L} \frac{\partial L}{\partial \theta} \right]^2 \geq 1. \quad (4.3)$$

Полагая $L \neq 0$, можно записать

$$\begin{aligned}\frac{\partial \ln L}{\partial \theta} &= \frac{1}{L} \frac{\partial L}{\partial \theta}, \\ \frac{\partial^2 \ln L}{\partial \theta^2} &= -\frac{1}{L^2} \left[\frac{\partial L}{\partial \theta} \right]^2 + \frac{1}{L} \frac{\partial^2 L}{\partial \theta^2}.\end{aligned}\quad (4.4)$$

Умножим (4.4) на L и получим математическое ожидание

$$M \left[\frac{1}{L} \frac{\partial^2 L}{\partial \theta^2} \right] = \int \dots \int \frac{\partial^2 L}{\partial \theta^2} dx_1, \dots, dx_n = 0$$

и

$$M \left[\frac{1}{L} \frac{\partial L}{\partial \theta} \right]^2 = -M \left[\frac{\partial^2 \ln L}{\partial \theta^2} \right]. \quad (4.5)$$

Учитывая (4.5) запишем неравенство (4.3) в виде

$$D(\hat{\theta}) = M(\hat{\theta} - \theta)^2 \geq \frac{1}{M \left[-\frac{\partial^2 \ln L}{\partial \theta^2} \right]}. \quad (4.6)$$

Неравенство (4.6) называется *неравенством Рао–Крамера*.

Выражение

$$M \left[-\frac{\partial^2 \ln L}{\partial \theta^2} \right] = M \left[\frac{\partial \ln L}{\partial \theta} \right]^2 = I^{(F)}(\theta)$$

называется информацией Фишера. Обратная величина от информации Фишера является нижней границей для дисперсии оценки $\hat{\theta}$

$$D(\hat{\theta}) \geq (I^{(F)}(\theta))^{-1}.$$

В случае вектора параметров $\vec{\theta} = (\theta_1, \dots, \theta_s)$, неравенство Рао–Крамера записывается следующим образом

$$D(\hat{\vec{\theta}}) \geq (I^{(F)}(\vec{\theta}))^{-1},$$

где

$$D_{ss'} = M[(\hat{\theta}_s - \theta_s)(\hat{\theta}_{s'} - \theta_{s'})]$$

– ковариационная матрица оценки вектора параметров и

$$I_{ss'}^{(F)} = M \left[-\frac{\partial^2 \ln L(\vec{x}, \vec{\theta})}{\partial \theta_s \partial \theta_{s'}} \right] = M \left[\frac{\partial \ln L(\vec{x}, \vec{\theta})}{\partial \theta_s} \frac{\partial \ln L(\vec{x}, \vec{\theta})}{\partial \theta_{s'}} \right]$$

– *информационная матрица Фишера*. Оценка $\hat{\theta}$ называется *эффективной*, если неравенство Рао–крамера переходит в равенство

$$D(\hat{\theta}) = [I^{(F)}(\theta)]^{-1}.$$

4.5 Достаточность

Если оценка $\hat{\theta}(x_1, \dots, x_n)$ содержит всю информацию о параметре θ , которая доступна из наблюдений (x_1, \dots, x_n) , говорят о *достаточной оценке*.

В случае повторной выборки плотность вероятностей $L(\vec{x}, \theta)$ при достаточном оценивании можно представить в виде:

$$L(\vec{x}, \theta) = g(\hat{\theta}, \theta)h(\vec{x}).$$

4.6 Робастное оценивание

Оценка называется *робастной*, если ее распределение нечувствительно к малым изменениям распределения выборочных величин или, в более общем случае, нечувствительно к малым отклонениям от используемых моделей.

Глава 5

Модели измерений

Первым шагом геофизической интерпретации является создание *модели измерений* (*модели экспериментального материала, математической модели экспериментального материала*). Такая модель является функциональным соотношением между экспериментальными наблюдениями (данными) и неизвестными параметрами геофизических объектов или среды. Вид функциональной связи определяется на основе физических законов с учетом особенностей проводимого эксперимента. Модели, используемые для оценки параметров должны учитывать случайный характер эксперимента. Свойства случайных компонент, присутствующих в модели, являются важной частью модели измерений.

5.1 Аддитивные модели

Пусть измерения геофизического поля u проводятся в точках x_k, y_l, z_m при значениях времени t_i :

$$u(x_k, y_l, z_m, t_i), \quad x_k = k\Delta x, \quad y_l = l\Delta y, \quad z_m = m\Delta z, \quad t_i = i\Delta t.$$

Пусть неизвестный вектор параметров геофизического поля:

$$\vec{\theta} = (\theta_1, \dots, \theta_s),$$

где s – число параметров. Пусть функциональное соотношение между наблюдаемыми величинами u и неизвестными параметрами $\vec{\theta}$ не описывается зависимостью $f(\vec{\theta}, x_k, y_l, z_m, t_i)$. Модель среды вида

$$u(x_k, y_l, z_m, t_i) = f(x_k, y_l, z_m, t_i, \vec{\theta}) + \varepsilon(x_k, y_l, z_m, t_i),$$

в которой случайная (шумовая) компонента является аддитивной, называется аддитивной моделью. Подобные модели наиболее часто используются при статистической интерпретации.

5.2 Модели количественной интерпретации

Аддитивная модель, рассмотренная ранее, описывает тип моделей, которые носят название *модели количественной интерпретации*. The additive model, introduced above, concerns to the type of models of the *quantitative interpretation*. задача количественной интерпретации состоит в оценивании вектора неизвестных параметров $\vec{\theta}$ с использованием методов математической статистики.

Пример 1: Модель сейсмической трассы:

$$u(t_i) = \sum_{\mu=1}^m A_{\mu} \varphi(t_i - \tau_{\mu}) + \varepsilon(t_i),$$

где $\varphi(t)$ – форма сейсмического сигнала, A_{μ} – амплитуда и τ_{μ} – время прихода сейсмического сигнала с номером μ . Вектор параметров в этом случае имеет вид The parameter vector has in this case the form

$$\vec{\theta} = \|A_{\mu}, \tau_{\mu}\|_{\mu=1}.$$

Пример сейсмической трассы представлен на рис. 5.1a. Сейсмическая трасса – результат численного решения The trace is a result of numerical simulation 1-D волнового уравнения для случая кусочно-од/-но/-род/-но/-го (рис. 5.1b) полупространства со свободной поверхностью. На сейсмической трассе можно выделить три волны, при этом каждая волна может быть описана амплитудой A_{μ} и временем прихода τ_{μ} . Непрерывную по времени модель можно

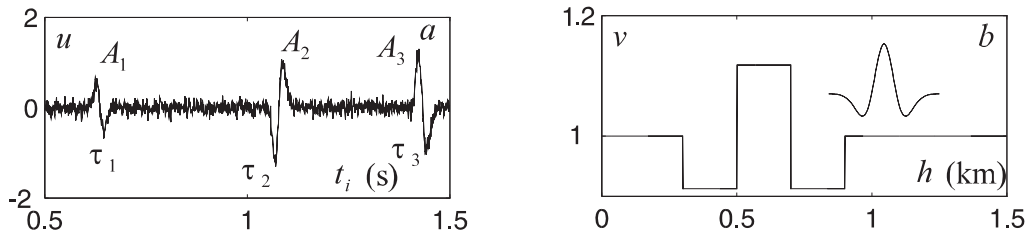


Рис. 5.1: Сейсмическая трасса. (a) – сейсмическая трасса осложненная нормальным шумом ($N(0, 0.1)$); (b) – скорость распространения волн в среде как функция глубины и временная зависимость воздействия в источнике. Источник расположен вблизи свободной поверхности.

представить в частотной области

$$u(\omega) = \sum_{\mu=1}^m A_{\mu} \Phi(\omega) \exp\{-i\omega\tau_{\mu}\} + E(\omega)$$

используя преобразование Фурье:

$$u(\omega) = \int_{-\infty}^{\infty} u(t) \exp\{-i\omega t\} dt, \quad \Phi(\omega) = \int_{-\infty}^{\infty} \varphi(t) \exp\{-i\omega t\} dt,$$

$$E(\omega) = \int_{-\infty}^{\infty} \varepsilon(t) \exp\{-i\omega t\} dt.$$

Дискретную временную модель, используя Z -преобразование ($Z = e^{-i\omega\Delta t}$) можно представить в виде

$$u(Z) = \sum_{\mu=1}^m A_{\mu} \Phi(Z) Z^{\tau_{\mu}/\Delta t} + E(Z).$$

Пример 2: Пусть вертикальная компонента магнитного поля, возбужденного источником типа магнитного диполя, зарегистрирована на линейном профиле, в этом случае модель измерений может быть записана в виде

$$u(x_k) = \frac{M(2h^2 - x_k^2)}{(x_k^2 + h^2)^{5/2}} + \varepsilon(x_k), \quad (5.1)$$

где M – магнитный момент диполя, h – глубина расположения диполя. Магнитное поле, представленное на рис. 5.2 удовлетворяет формуле (5.1) ($h = 1$ км, $M = 1$).

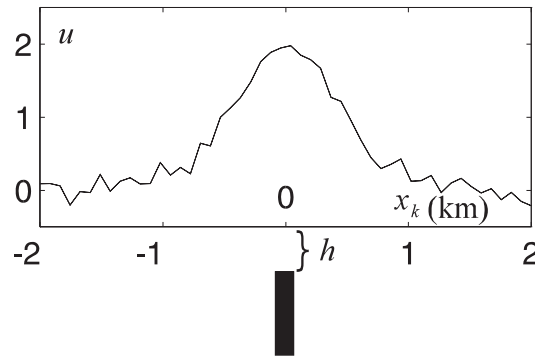


Рис. 5.2: Магнитное поле осложненное нормальным шумом ($N(0, \sigma)$). Магнитное поле, отвечающее регистрации на линейном профиле, вычислено по формуле (5.1) ($h = 1$ км, $M = 1$, $\sigma = 0.1$).

5.3 Регрессионная модель

Для аппроксимации геофизических наблюдений очень часто используется *регрессионная* модель или линейная модель. Линейное соотношение между искомыми параметрами и наблюдениями, в общем случае, может достигаться путем линеаризации используемых физических связей и математических выражений

$$u(x_k, y_l, z_m, t_i) = \sum_{j=0}^S \theta_j \psi_j(x_k, y_l, z_m, t_i) + \varepsilon(x_k, y_l, z_m, t_i).$$

Пример 1: В случае обработки двумерного магнитного поля u , мы можем записать модель в виде:

$$u(x_k, z_m) = \sum_{j=0}^S \theta_j \psi_j(x_k, z_m) + \varepsilon(x_k, z_m),$$

где θ_j – магнитный момент аномалии с номером j .

Пример 2: Модель сейсмограммы:

$$u(x_k, t_i) = \sum_{j=0}^S \theta_j \psi_j(x_k, t_i) + \varepsilon(x_k, t_i),$$

где θ_j – амплитуда сейсмической волны, $\psi_j(x_k, t_i)$ – форма сейсмической волны с неизвестной временной задержкой.

Очень часто на практике для аппроксимации геофизических данных используются одномерные модели

Пример 3: Наблюдение магнитного поля вдоль линейного профиля:

$$u(x_k) = \sum_{j=0}^S \theta_j \psi_j(x_k) + \varepsilon(x_k).$$

Пример 4: Сейсмическая трасса, зарегистрированная одним геофоном:

$$u(t_i) = \sum_{j=0}^S \theta_j \psi_j(t_i) + \varepsilon(t_i).$$

Рассмотренную ранее линейную модель можно записать в компактной форме с использованием векторных и матричных обозначений:

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon},$$

где

$$\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix}, \quad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_k \end{pmatrix}, \quad \psi = \begin{pmatrix} \psi_{11} & \dots & \psi_{1S} \\ \dots & \dots & \dots \\ \psi_{K1} & \dots & \psi_{KS} \end{pmatrix}$$

Матрицу ψ называют *матрицей плана* или *структурной матрицей*.

5.4 Модели количественной интерпретации

Если в результате интерпретации данных возможно получить информацию относительно качественного состояния объекта, то соответствующая модель измерений называется моделью *качественной интерпретации*.

$$u(x_k, y_l, z_m, t_i) = \begin{cases} \text{either} & f_1(x_k, y_l, z_m, t_i) + \varepsilon(x_k, y_l, z_m, t_i), \\ \text{or} & f_2(x_k, y_l, z_m, t_i) + \varepsilon(x_k, y_l, z_m, t_i)' \\ \dots & \dots \\ \text{or} & f_N(x_k, y_l, z_m, t_i) + \varepsilon(x_k, y_l, z_m, t_i)' \end{cases}$$

$$\nu = 1, 2, \dots, N$$

где f_ν – поле объекта с состояниями, пронумерованными $\nu = 1, 2, \dots, N$.

Пример 1: Обнаружение сейсмического сигнала на фоне шума:

$$u(t_i) = \begin{cases} \text{either} & A\varphi(t_i - \tau) + \varepsilon(t_i), \\ \text{or} & \varepsilon(t_i), \end{cases}$$

где A – известная амплитуда, τ – известное время прихода сейсмического сигнала (см. рис. 5.3). Задача состоит в обнаружении сейсмического сигнала или в фиксации отсутствия сигнала.

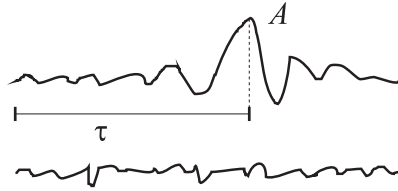


Рис. 5.3: Качественная интерпретация (“шум и сигнал или шум”).

Пример 2: Выбор между двумя сейсмическими сигналами

$$u(t_i) = \begin{cases} \text{either} & A_1\varphi(t_i - \tau_1) + \varepsilon(t_i) \\ \text{or} & A_2\varphi(t_i - \tau_2) + \varepsilon(t_i) \end{cases}$$

где A_1 и τ_1 – амплитуда и время прихода 1-й волны, A_2 и τ_2 – параметры 2-й волны. Задача состоит в обработке сейсмической трассы с целью определения принадлежности сигнала к одному из двух заданных (см. рис. 5.4).

Пример 3: Разрешение сейсмических сигналов во временной области:

$$u(t_i) = \begin{cases} \text{either} & A_0\varphi_0(t_i - \tau_0) + \varepsilon(t_i) \\ \text{or} & A_1\varphi_1(t_i - \tau_1) + A_2\varphi_2(t_i - \tau_2) + \varepsilon(t_i). \end{cases}$$

Задача состоит в принятии решения: на интервале сейсмической трассы имеется одна волна с параметрами (A_0, τ_0) или две волны с параметрами (A_1, τ_1) and (A_2, τ_2) (см. рис. 5.5).

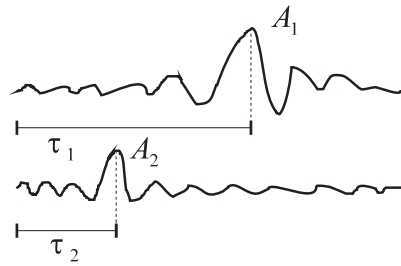


Рис. 5.4: Качественная интерпретация (“сигнал (номер 1) и шум или сигнал (номер 2) и шум”).

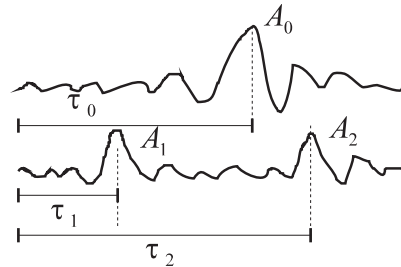


Рис. 5.5: Качественная интерпретация (“сигнал и шум или два сигнала и шум”).

Пример 4: Разрешение сейсмических сигналов в частотной области (см. рис. 5.6):

$$u(\omega) = \begin{cases} \text{either} & A_0\Phi_0(\omega - \omega_0) + E(\omega) \\ \text{or} & A_1\Phi_1(\omega - \omega_1) + A_2\Phi_2(\omega - \omega_2) + E(\omega). \end{cases}$$

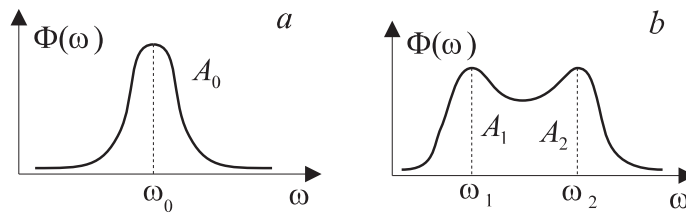


Рис. 5.6: Качественная интерпретация в частотной области (“сигнал и шум или два сигнала и шум”).

5.5 Модели качественно-количественной интерпретации

При решении сложных интерпретационных задач, геофизик имеет дело с моделями следующего типа:

$$u(x_k, y_l, z_m, t_i) = \begin{cases} \text{either} & f_1(x_k, y_l, z_m, t_i, \vec{\theta}_1) + \varepsilon(x_k, y_l, z_m, t_i), \\ \text{or} & f_2(x_k, y_l, z_m, t_i, \vec{\theta}_2) + \varepsilon(x_k, y_l, z_m, t_i), \\ \dots & \dots \\ \text{or} & f_N(x_k, y_l, z_m, t_i, \vec{\theta}_N) + \varepsilon(x_k, y_l, z_m, t_i), \end{cases}$$

$$\nu = 1, 2, \dots, N$$

где состояние исследуемого объекта характеризуется N различными неслучайными состояниями или сигналами f_ν , при этом каждое состояние зависит от вектора $\vec{\theta}_\nu$ неизвестных параметров. Задача состоит в выборе оптимальной процедуры выбора между возможными состояниями и в оценивании неизвестных параметров. Следует отметить, что при создании модели необходимо задать (определить) свойства случайной компоненты.

5.6 Случайная компонента модели и ее свойства

Структура алгоритма оценивания определяется свойствами случайной компоненты модели, поэтому необходим тщательный анализ свойств случайной компоненты.

5.6.1 Нормальное распределение

Принимая во внимание физическую природу шума, с учетом центральной предельной теоремы мы можем часто полагать, что распределение шума близко к *нормальному распределению*

$$\vec{\varepsilon} \in N(\vec{\varepsilon}_0, R).$$

Плотность вероятностей случайного (нормально распределенного) вектора $\vec{\varepsilon}$:

$$p(\vec{\varepsilon}) = (2\pi)^{-n/2} |R|^{-1/2} \exp\left\{-\frac{1}{2}(\vec{\varepsilon} - \vec{\varepsilon}_0)^T R^{-1}(\vec{\varepsilon} - \vec{\varepsilon}_0)\right\},$$

где элементы ковариационной матрицы определяются следующим образом

$$R_{ii'} = M[(\varepsilon_i - \varepsilon_{0i})(\varepsilon_{i'} - \varepsilon_{0i'})]$$

и элементы корреляционной матрицы имеют вид

$$r_{ii'} = \frac{R_{ii'}}{\sqrt{R_{ii}}\sqrt{R_{i'i'}}}.$$

5.6.2 Случайная компонента как стационарный процесс

Корреляционную матрицу можно записать в виде:

$$r = \begin{bmatrix} 1 & r_1 & r_2 & \dots & r_n \\ \dots & \dots & \dots & \dots & \dots \\ r_n & r_{n-1} & r_{n-2} & \dots & 1 \end{bmatrix}.$$

5.6.3 Случайная компонента как некоррелированный процесс

Элементы ковариационной матрицы можно записать в виде:

$$R_{ii'} = \sigma_i^2 \delta_{ii'}, \quad \delta_{ii'} = \begin{cases} 1 & \text{by } i = i', \\ 0 & \text{by } i \neq i', \end{cases}$$

$$R = \sigma^2 E, \quad R^{-1} = \frac{1}{\sigma^2} E,$$

где E – единичная матрица. Нормальная плотность имеет вид:

$$p(\vec{\varepsilon}) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}(\vec{\varepsilon} - \vec{\varepsilon}_0)^2\right\}. \quad (5.2)$$

5.6.4 Случайная компонента как марковский процесс

Элементы ковариационной матрицы в этом случае можно записать в виде

$$R_{ii'} = \sigma^2 \exp\{-c|i - i'|\} \quad (5.3)$$

(см. рис. 5.7). Вводя обозначение

$$\gamma = \exp(-c)$$

мы можем записать ковариационную матрицу:

$$R = \sigma^2 \begin{bmatrix} 1 & \gamma & \gamma^2 & \dots & \gamma^n \\ \gamma & 1 & \gamma & \dots & \gamma^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma^n & \gamma^{n-1} & \gamma^{n-2} & \dots & 1 \end{bmatrix}$$

Определитель R равен

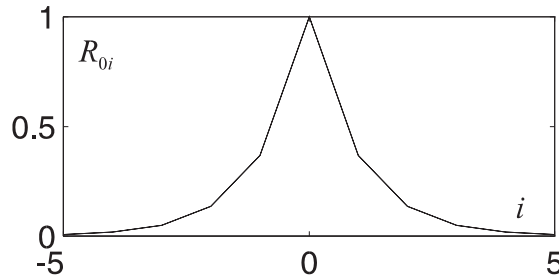


Рис. 5.7: Элементы ковариационной матрицы в случае марковского процесса (5.3) ($\sigma = 1, c = 1$).

$$|R| = \sigma^{2n}(1 - \gamma^2)^{n-1}.$$

Обратная матрица имеет вид

$$R^{-1} = \frac{1}{\sigma^2(1-\gamma^2)} \begin{bmatrix} 1 & -\gamma & 0 & \dots & 0 \\ -\gamma & 1+\gamma^2 & -\gamma & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Нормальная плотность для марковского процесса:

$$p(\vec{\varepsilon}) = (2\pi\sigma^2)^{-n/2}(1-\gamma^2)^{(n-1)/2} \exp\left\{-\frac{1}{2\sigma^2(1-\gamma^2)} \times \right. \\ \left. \times [(1+\gamma^2) \sum_{i=1}^n \varepsilon_i^2 - 2\gamma \sum_{i=1}^n \varepsilon_i \varepsilon_{i+1} - \gamma^2(\varepsilon_1^2 + \varepsilon_n^2)]\right\}.$$

5.6.5 Распределение Лапласа

Распределение Лапласа

$$p(\varepsilon) = \frac{1}{2\alpha} \exp\left\{-\frac{|\varepsilon - \varepsilon_0|}{\alpha}\right\}$$

является более эффективным, чем нормальное распределение для описания случайной компоненты при наличии выбросов. “Хвосты” распределения Лапласа спадают медленнее, чем “хвосты” нормального распределения (5.2) (see Figure 5.8).

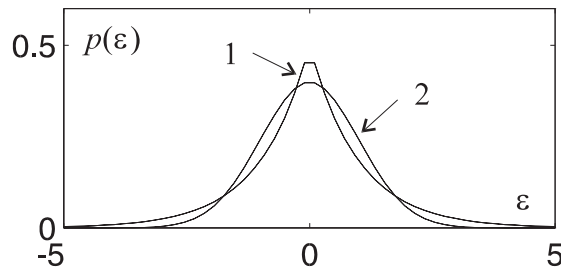


Рис. 5.8: Распределение Лапласа и нормальное распределение. 1 – распределение Лапласа ($\varepsilon_0 = 0$, $\alpha = 1$); 2 – нормальное распределение ($\varepsilon_0 = 0$, $\sigma = 1$).

5.7 Модель со случайным параметром

Пусть параметр θ – случайная величина. Запишем модель в виде

$$u(x_k, t_i) = \sum_{\mu=1}^M A_{\mu} \varphi(t_i - \tau_{\mu k}) + \varepsilon(x_k, t_i),$$

где A_μ и τ_μ – соответственно случайная амплитуда и случайное время прихода волны с номером μ . Форма волны $\varphi(t_i)$ не является случайной. При решении геофизических проблем (на практике) часто предполагают, что случайный параметр распределен по нормальному закону.

$$A \in N(0, \sigma_A^2), \quad \tau \in N(0, \sigma_\tau^2), \quad \varepsilon \in N(0, \sigma_\varepsilon^2).$$

5.8 Априорная информация о поле параметров

Под априорной информацией понимаю информацию не связанную с результатами проводимого эксперимента.

Априорная информация, например, в виде системы равенств

$$\vec{\Phi}(\vec{\theta}) = \vec{c}$$

ли неравенств

$$\vec{c}_1 \leq \vec{\Phi}(\vec{\theta}) \leq \vec{c}_2,$$

где $\vec{\theta}$ – вектор неизвестных параметров, $\vec{\Phi}(\vec{\theta})$ – заданная векторная функция.

Для практического применения априорную информацию более удобно задавать в виде системы линейных равенств

$$A\vec{\theta} = \vec{V}$$

или неравенств

$$\vec{V}_1 \leq A\vec{\theta} \leq \vec{V}_2,$$

где \vec{V}_1 и \vec{V}_2 – соответственно заданные нижняя и верхняя границы.

Если априорная информация относительно неизвестного параметра θ отсутствует, то он (формально) может принимать значения от $-\infty$ до $+\infty$. Следовательно такому случаю отвечает плотность вероятностей вида

$$p(\theta) = \text{const} \quad \text{with} \quad -\infty < \theta < \infty.$$

Такая плотность вероятностей не является плотностью вероятностью “в строгом смысле”, т.к. $\int_{-\infty}^{\infty} p(\theta) d\theta \neq 1$.

Ранее отмечалось, что энтропия является мерой неопределенности системы. Применяя принцип максимальной энтропии можно брать распределение, которое доставляет максимальную неопределенность искомым величинам при заданных ограничениях. Предположение об отсутствии априорной информации может быть трансформировано в предположение о наличии априорной информации содержащей максимальную возможную неопределенность относительно искомым параметров.

Рассмотрим принцип максимальной энтропии для определения распределений с максимальной неопределенностью при заданных ограничениях:

1. Плотность вероятностей случайного вектора $\vec{\theta} = \{\theta_1, \dots, \theta_s\}$ с компонентами, определенными на интервалах $(\theta'_s - \theta''_s)$ обладает максимальной неопределенностью (максимальной энтропией) в случае *равномерного распределения*

$$p(\vec{\theta}) = \begin{cases} \prod_{s=1}^S 1/(\theta'_s - \theta''_s) & \theta''_s \leq \theta_s \leq \theta'_s, \\ 0 & \theta_s. \end{cases}$$

2. Рассмотрим плотность вероятностей случайного вектора $\vec{\theta} = \{\theta_1, \dots, \theta_s\}$ с математическим ожиданием $\langle \vec{\theta} \rangle$ и ковариационной матрицей $\sigma_{\theta_s}^2$ с компонентами, принадлежащими интервалу $(-\infty, \infty)$. В этом случае максимально энтропийным распределением является *нормальное распределение*

$$p(\vec{\theta}) = \prod_{s=1}^S (2\pi\sigma_{\theta_s}^2)^{-1/2} \exp\left\{-\frac{1}{2} \sum_{s=1}^S \frac{(\theta_s - \langle \theta_s \rangle)^2}{\sigma_{\theta_s}^2}\right\}.$$

Пример: Рассмотрим оценивание параметров сейсмического сигнала

$$f(A, \tau) = A\varphi(t_i - \tau).$$

Пусть амплитуда A имеет нормальное распределение с математическим ожиданием $\langle A \rangle$ и дисперсией σ_A^2 . Время прихода сейсмического сигнала τ имеет равномерное распределение. Такую априорную информацию можно представить в виде

$$p(A, \tau) = \begin{cases} [\sqrt{2\pi}\sigma_A(\tau' - \tau'')]^{-1} \times & \text{by } \tau'' \leq \tau \leq \tau', \\ \times \exp\{-(A - \langle A \rangle)^2/2\sigma_A^2\} & \\ 0 & \text{for others } \tau. \end{cases}$$

Глава 6

Решение обратных геофизических задач: линейные модели

Пусть геофизические наблюдения, несущие информацию о неизвестном векторе параметров $\vec{\theta}$ представлены в виде вектора \vec{u} . Неизвестные параметры $\vec{\theta}$ будут оцениваться с использованием решающей функции $\lambda(\vec{u}, \vec{\theta})$. Рассмотрим линейные модели измерений.

$$\vec{u} = \psi\vec{\theta} + \vec{\varepsilon},$$

при этом начнем с метода наименьших квадратов.

6.1 Метод наименьших квадратов (МНК)

Метод наименьших квадратов при оценивании неизвестных параметров $\vec{\theta}$ заключается в минимизации суммы взвешенных квадратов

$$\begin{aligned}\lambda(\vec{\theta}) &= (\vec{u} - \psi\vec{\theta})^T W (\vec{u} - \psi\vec{\theta}), \\ \hat{\vec{\theta}} &= \arg \min \lambda(\vec{\theta}),\end{aligned}$$

где W – диагональная весовая функция $W = \text{diag}(w_1, \dots, w_n)$. Для нахождения минимума, функция λ дифференцируется по каждому из искомым параметрам и и каждое из получаемых при этом уравнений приравнивается нулю

$$\partial\lambda/\partial\theta_s = 0 \quad \Rightarrow \quad (\psi^T W \psi)\vec{\theta} = \psi^T W \vec{u}.$$

Мы можем интерпретировать весовую матрицу как обратную ковариационную матрицу случайной компоненты $\vec{\varepsilon}$

$$W = R_\varepsilon^{-1} \quad \text{or} \quad W = \sigma_\varepsilon^{-2} I,$$

где σ_ε^{-2} – дисперсия случайной компоненты.

$$\hat{\vec{\theta}} = (\psi^T \psi)^{-1} \psi^T \vec{u}.$$

Ковариационную матрицу оценки $\hat{\vec{\theta}}$ можно представить в виде

$$\begin{aligned} R_{\theta} &= M[(\hat{\vec{\theta}} - \vec{\theta})(\hat{\vec{\theta}} - \vec{\theta})^T] = M[(\psi^T \psi)^{-1} \psi^T \vec{u} - \vec{\theta}][\vec{u}^T \psi (\psi^T \psi)^{-1} - \vec{\theta}^T] = \\ &= M[(\psi^T \psi)^{-1} (\psi^T \psi) \vec{\theta} - \vec{\theta} + (\psi^T \psi)^{-1} \psi^T \vec{\varepsilon}][(\psi^T \psi) (\psi^T \psi)^{-1} \vec{\theta}^T - \vec{\theta}^T + \vec{\varepsilon}^T \psi (\psi^T \psi)^{-1}] \\ &= (\psi^T \psi)^{-1} \psi^T M[\vec{\varepsilon} \vec{\varepsilon}^T] \psi (\psi^T \psi)^{-1} = \sigma_{\varepsilon}^2 (\psi^T \psi)^{-1}. \end{aligned}$$

6.1.1 Свойства МНК оценок

1. Оценка $\hat{\theta}_{MLS}$ является несмещенной:

$$\begin{aligned} M[\hat{\vec{\theta}}] &= (\psi^T \psi)^{-1} \psi^T M[\vec{u}] = \\ &= (\psi^T \psi)^{-1} \psi^T M[\psi \vec{\theta} + \varepsilon] = \vec{\theta}, \end{aligned}$$

что следует из предположения $M[\varepsilon] = 0$.

2. Оценка $\hat{\theta}_{MLS}$ является эффективной:

$$R_{ss}(\hat{\vec{\theta}}) \leq R_{ss}(\vec{\theta}^*) \quad s = 1, \dots, S,$$

где $\vec{\theta}^*$ – произвольная оценка. Оценка по МНК имеет минимальную оценку в классе линейных оценок.

3. Среднеквадратическая ошибка вектора измерений \vec{u} и его оценки $\hat{\vec{u}} = \psi \hat{\vec{\theta}}$ связаны соотношениями

$$\begin{aligned} \hat{\sigma}_{\varepsilon}^2 &\Rightarrow M[(\vec{u} - \psi \hat{\vec{\theta}})^T (\vec{u} - \psi \hat{\vec{\theta}})] = \sigma_{\varepsilon}^2 (n - S) \Rightarrow \\ &\Rightarrow \hat{\sigma}_{\varepsilon}^2 = (\vec{u} - \psi \hat{\vec{\theta}})^T (\vec{u} - \psi \hat{\vec{\theta}}) / (n - S) = \hat{\varepsilon}^T \hat{\varepsilon} / (n - S). \end{aligned}$$

Пример: Найдем амплитуды сейсмических волн с известными временами прихода s τ_s .

Модель сейсмической трассы:

$$\vec{u} = \psi \vec{A} + \vec{\varepsilon},$$

где $\psi = \|\varphi(t_i - \tau_s)\|_{n \times S}$ – матрица, $\vec{\varepsilon}$ – некоррелированный случайный шум с нормальным распределением. Оценка амплитуд по МНК имеет вид

$$\hat{\vec{A}} = (\psi^T \psi)^{-1} \psi^T \vec{u}$$

с ковариационной матрицей оценки $\hat{\vec{A}}$

$$R_{\hat{A}} = \hat{\sigma}_{\varepsilon}^2 (\psi^T \psi)^{-1},$$

где

$$\hat{\sigma}_{\varepsilon}^2 = \frac{(\vec{u} - \psi \hat{\vec{A}})^T (\vec{u} - \psi \hat{\vec{A}})}{n - S}$$

– среднеквадратическая ошибка.

6.2 Метод наименьших квадратов: нелинейная модель

В общем случае связь между параметрами $\vec{\theta}$ и наблюдениями \vec{u} не является линейной

$$\vec{u} = \vec{\varphi}(\vec{\theta}) + \vec{\varepsilon},$$

где $\varphi_i(\vec{\theta})$ – дифференцируемая функция неизвестных параметров.

Мы можем линеаризовать функции $\varphi_i(\vec{\theta})$ воспользовавшись разложением в ряд Тейлора. Считая известными (опорные) параметры $\vec{\theta}_0$ получаем:

$$\vec{\varphi}(\vec{\theta}) = \vec{\varphi}(\vec{\theta}_0) + \psi \Delta\vec{\theta} + \dots$$

Модель можно переписать в виде

$$\tilde{\vec{u}} = \psi \Delta\vec{\theta} + \vec{\varepsilon},$$

где $\tilde{\vec{u}} = \vec{u} - \vec{\varphi}(\vec{\theta}_0)$, ψ – матрица,

$$\|\psi_{js}\| = \partial\varphi_j / \partial\theta_s |_{\vec{\theta}=\vec{\theta}_0}$$

и $\Delta\vec{\theta} = \vec{\theta} - \vec{\theta}_0$.

Оценки параметров используются для итеративного вычисления новых оценок.

Оценка, полученная на первом шаге имеет вид:

$$\Delta\hat{\vec{\theta}}^{(1)} = (\psi_0^T W_0 \psi_0)^{-1} \psi_0^T W_0 \tilde{\vec{u}}_0, \quad \hat{\vec{\theta}}^{(1)} = \vec{\theta}_0 + \Delta\hat{\vec{\theta}}^{(1)},$$

где W_0 – диагональная весовая матрица

$$W_0 = (\hat{\sigma}_{\varepsilon 0}^2)^{-1} I,$$

при этом среднеквадратическая ошибка:

$$\hat{\sigma}_{\varepsilon 0}^2 = \frac{(\vec{u} - \psi_0 \vec{\theta}_0)^T (\vec{u} - \psi_0 \vec{\theta}_0)}{n - S}.$$

Оценка на втором шаге:

$$\Delta\hat{\vec{\theta}}^{(2)} = (\psi_1^T W_1 \psi_1)^{-1} \psi_1^T W_1 \tilde{\vec{u}}_1,$$

где $\tilde{\vec{u}}_1 = \vec{u} - \psi_1 \hat{\vec{\theta}}^{(1)}$ – новый вектор измерений, матрица ψ_1 :

$$\|\psi_{js}^{(1)}\| = \frac{\partial\varphi_j}{\partial\theta_s} \Big|_{\vec{\theta}=\hat{\vec{\theta}}_1}$$

и весовая матрица на втором шаге:

$$W_1 = (\hat{\sigma}_{\varepsilon_1}^2)^{-1} I = \left[\frac{(\vec{u} - \psi_0 \vec{\theta}_1)^T (\vec{u} - \psi_0 \vec{\theta}_1)}{n - S} \right]^{-1} I.$$

Проверка выполнения порогового условия. Если неравенство выполняется:

$$\left| \frac{\Delta \hat{\theta}_s^{(2)}}{\hat{\theta}_s^{(2)}} \right| \leq \beta,$$

то итерационный процесс прекращается. В альтернативном случае итерационный процесс продолжается.

Например, мы можем использовать такую процедуру для оценивания времен прихода τ_s сигналов в модели

$$\vec{u} = \psi \vec{A} + \vec{\varepsilon},$$

(см. 6.1) где $\psi = \|\varphi(t_i - \tau_s)\|$ – матрица ($n \times S$).

6.3 Метод наименьших квадратов: ортогональные полиномы

Полиномиальные модели часто используются для сглаживания кривых или поверхностей. Например, в случае линейной модели

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon}$$

элементы матрицы ψ могут иметь вид $\psi_{is} = x_i^s$ – очень важный специальный случай.

Мы можем записать i -ю строку матрицы ψ как

$$\psi_i = \|\varphi_0(x_i), \varphi_1(x_i), \dots, \varphi_{S-1}(x_i)\|.$$

Условие ортогональности:

$$\sum_{i=1}^n \varphi_s(x_i) \varphi_{s'}(x_i) = \begin{cases} \sum_{i=1}^n \varphi_s^2(x_i) & \text{for } s = s' \\ 0 & \text{for } s \neq s' \end{cases}$$

и $\psi^T \psi$ записывается в виде

$$\psi^T \psi = \left\| \begin{array}{cccc} \sum_{i=1}^n \varphi_0^2(x_i) & 0 & \dots & 0 \\ 0 & \sum_{i=1}^n \varphi_1^2(x_i) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sum_{i=1}^n \varphi_{S-1}^2(x_i) \end{array} \right\|.$$

В специальном случае полиномов Чебышева:

$$\varphi_0(x_i) = 1, \quad \varphi_1(x_i) = x_i + b_0 \varphi_0(x_i).$$

Для нахождения коэффициента b_0 выполняем умножение на $\varphi_0(x_i)$ и суммируем

$$b_0 \Rightarrow \sum_{i=1}^n \varphi_1(x_i) \varphi_0(x_i) = \sum_{i=1}^n x_i \varphi_0(x_i) + b_0 \sum_{i=1}^n \varphi_0^2(x_i).$$

Учитывая, что

$$\sum_{i=1}^n \varphi_1(x_i) \varphi_0(x_i) = 0,$$

получаем

$$b_0 = -\frac{\sum_{i=1}^n x_i \varphi_0(x_i)}{\sum_{i=1}^n \varphi_0^2(x_i)},$$

и можем записать $\varphi_1(x_i)$ в явном виде

$$\varphi_1(x_i) = x_i - \frac{\sum_{i=1}^n x_i \varphi_0(x_i)}{\sum_{i=1}^n \varphi_0^2(x_i)} \varphi_0(x_i).$$

Для функции $\varphi_2(x_i)$ имеем

$$\varphi_2(x_i) = x_i^2 + b_1 \varphi_1(x_i) + b_0 \varphi_0(x_i).$$

Коэффициенты b_0 и b_1 находим используя условия ортогональности

$$\begin{aligned} \sum_{i=1}^n \varphi_2(x_i) \varphi_0(x_i) = 0, \quad b_0 &= -\frac{\sum_{i=1}^n x_i^2 \varphi_0(x_i)}{\sum_{i=1}^n \varphi_0^2(x_i)}, \\ \sum_{i=1}^n \varphi_2(x_i) \varphi_1(x_i) = 0, \quad b_1 &= -\frac{\sum_{i=1}^n x_i^2 \varphi_1(x_i)}{\sum_{i=1}^n \varphi_1^2(x_i)}. \end{aligned}$$

Подставляя выражения для b_0 и b_1 в функцию $\varphi_2(x_i)$, получаем

$$\varphi_2(x_i) = x_i^2 - \frac{\sum_{i=1}^n x_i^2 \varphi_1(x_i)}{\sum_{i=1}^n \varphi_1^2(x_i)} \varphi_1(x_i) - \frac{\sum_{i=1}^n x_i^2 \varphi_0(x_i)}{\sum_{i=1}^n \varphi_0^2(x_i)} \varphi_0(x_i).$$

Запишем функцию $\varphi_s(x_i)$ в виде

$$\varphi_s(x_i) = x_i^s - \frac{\sum_{i=1}^n x_i^s \varphi_{s-1}(x_i)}{\sum_{i=1}^n \varphi_{s-1}^2(x_i)} \varphi_{s-1}(x_i) - \dots - \frac{\sum_{i=1}^n x_i^s \varphi_0(x_i)}{\sum_{i=1}^n \varphi_0^2(x_i)} \varphi_0(x_i),$$

$$\varphi_s(x_i) = x_i^s + b_{s-1} \varphi_{s-1}(x_i) + \dots + b_0 \varphi_0(x_i).$$

Коэффициенты b_q ($q < s$) определяются с использованием условий ортогональности

$$\sum_{i=1}^n x_i^s \varphi_q(x_i) + b_q \sum_{i=1}^n \varphi_q^2(x_i) = 0, \quad b_q = -\frac{\sum_{i=1}^n x_i^s \varphi_q(x_i)}{\sum_{i=1}^n \varphi_q^2(x_i)}.$$

Оценки параметров $\hat{\theta}_s$ определяются как

$$\hat{\theta}_s = \sum_{j=1}^n \varphi_{sj} u_j, \quad s = 0, 1, 2, \dots, S-1,$$

$$\hat{\theta} = \Phi \vec{u},$$

где матрица Φ имеет вид

$$\Phi = \begin{pmatrix} \varphi_0(x_1) & \varphi_0(x_2) & \dots & \varphi_0(x_n) \\ \varphi_1(x_1) & \varphi_1(x_2) & \dots & \varphi_1(x_n) \\ \dots & \dots & \dots & \dots \\ \varphi_{S-1}(x_1) & \varphi_{S-1}(x_2) & \dots & \varphi_{S-1}(x_n) \end{pmatrix}.$$

При выполнении условия ортонормальности:

$$\sum_{i=1}^n \varphi_{si}^2 = 1.$$

Пусть ковариационная матрица случайной компоненты $R_\varepsilon = \sigma_\varepsilon^2 I$, в этом случае получаем

$$R_\theta = \sigma_\varepsilon^2 I.$$

Оценка МНК:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n-S} (\vec{u} - \psi \hat{\theta})^T (\vec{u} - \psi \hat{\theta}).$$

6.4 Метод наименьших квадратов с ограничениями

Пусть на неизвестные параметры линейной модели

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon}$$

наложены ограничения

$$A\vec{\theta} = \vec{V},$$

где A обозначает $[K \times S]$ матрицу известных коэффициентов, и V – известный $[K \times 1]$ вектор.

Наилучшая несмещенная оценка $\tilde{\theta}$ вектора θ неизвестных параметров будет получена с использованием метода неопределенных множителей Лагранжа.

Построим функцию

$$\Phi(\vec{\theta}) = (\vec{u} - \psi\vec{\theta})^T(\vec{u} - \psi\vec{\theta}) - 2\vec{\lambda}^T(A\vec{\theta} - \vec{V}),$$

где $\vec{\lambda}$ – $(K \times 1)$ вектор множителей Лагранжа. Для нахождения минимума этой функции выполним дифференцирование по параметрам θ_s

$$\begin{aligned} \partial\Phi/\partial\theta_s &= 0, \quad s = 1, \dots, S, \\ -2\vec{u}^T\psi + 2\vec{\theta}^T\psi^T\psi - 2\vec{\lambda}^T A &= 0 \mid (\psi^T\psi)^{-1}. \end{aligned}$$

После умножения на матрицу $(\psi^T\psi)^{-1}$ и использования выражения для МНК оценок

$$\hat{\theta} = (\psi^T\psi)^{-1}\psi^T\vec{u}, \text{ получаем}$$

$$\tilde{\theta}^T = \hat{\theta}^T + \vec{\lambda}^T A(\psi^T\psi)^{-1} \mid A^T.$$

После умножения на матрицу A^T имеем

$$\tilde{\theta}^T A^T = \hat{\theta}^T A^T + \vec{\lambda}^T A(\psi^T\psi)^{-1} A^T.$$

Учитывая, что оценка $\tilde{\theta}$ должна удовлетворять линейным ограничениям

$$\tilde{\theta}^T A^T = V^T,$$

найдем вектор (неопределенных) множителей

$$\vec{\lambda}^T = (\vec{V}^T - \hat{\theta}^T A^T)(A(\psi^T\psi)^{-1} A^T)^{-1}$$

и, окончательно, получим оценку

$$\tilde{\theta}^T = \hat{\theta}^T + (\vec{V}^T - \hat{\theta}^T A^T)(A(\psi^T\psi)^{-1} A^T)^{-1} A(\psi^T\psi)^{-1}.$$

Как будет показано далее, ковариационная матрица $R_{\tilde{\theta}}$ оценки $\tilde{\theta}$ может быть записана в виде

$$R_{\tilde{\theta}} = R_{\hat{\theta}} - \sigma_{\varepsilon}^2(\psi^T\psi)^{-1} A^T (A(\psi^T\psi)^{-1} A^T)^{-1} A(\psi^T\psi)^{-1},$$

где

$$R_{\hat{\theta}} = \hat{\sigma}_\varepsilon^2 (\psi^T \psi)^{-1}, \quad \hat{\sigma}_\varepsilon^2 = \frac{1}{n - S - K} (\vec{u} - \psi \hat{\theta})^T (\vec{u} - \psi \hat{\theta}).$$

Запишем определение ковариационной матрицы $R_{\hat{\theta}}$

$$\begin{aligned} R_{\hat{\theta}} &= M[(\hat{\theta} - \vec{\theta})(\hat{\theta} - \vec{\theta})^T] = M[(\hat{\theta} - \vec{\theta} + (\psi^T \psi)^{-1} A^T \times \\ &\times (A(\psi^T \psi)^{-1} A^T)^{-1} (\vec{V} - A\hat{\theta}))(\hat{\theta}^T - \vec{\theta}^T + (\vec{V}^T - \hat{\theta}^T A^T) \times \\ &\times (A(\psi^T \psi)^{-1} A^T)^{-1} A(\psi^T \psi)^{-1})], \end{aligned}$$

вводя обозначения $B = (\psi^T \psi)^{-1}$, $C = (A(\psi^T \psi)^{-1} A^T)^{-1}$, получаем

$$R_{\hat{\theta}} = [(\hat{\theta} - \vec{\theta} + BA^T C(\vec{V} - A\hat{\theta})) \times (\hat{\theta}^T - \vec{\theta}^T + (\vec{V}^T - \hat{\theta}^T A^T) CAB)]$$

Учитывая, что

$$\begin{aligned} R_{\hat{\theta}} &= M[(\hat{\theta} - \vec{\theta})(\hat{\theta} - \vec{\theta})^T] = M[(B\psi^T(\psi\vec{\theta} + \vec{\varepsilon}) - \vec{\theta}) \times \\ &\times (B\psi^T(\psi\vec{\theta} + \vec{\varepsilon}) - \vec{\theta})^T] = BM[\vec{\varepsilon}\vec{\varepsilon}^T] = \sigma_\varepsilon^2 B \end{aligned}$$

и

$$\begin{aligned} &M[(\vec{V} - AB\psi^T(\psi\vec{\theta} + \vec{\varepsilon}))(B\psi^T(\psi\vec{\theta} + \vec{\varepsilon}) - \vec{\theta})^T] = \\ &= M[(\vec{V} - A\vec{\theta} - AB\psi^T\vec{\varepsilon})(\vec{\theta} + B\psi^T\vec{\varepsilon} - \vec{\theta})^T] = \\ &= AB\psi^T M[\vec{\varepsilon}\vec{\varepsilon}^T] \psi B = \sigma_\varepsilon^2 AB \end{aligned}$$

и

$$\begin{aligned} &M[(\vec{V} - AB\psi^T(\psi\vec{\theta} + \vec{\varepsilon}))(\vec{V}^T - (\vec{\varepsilon}^T + \vec{\theta}^T \psi^T) \psi BA^T)] = \\ &= AB\psi^T M[\vec{\varepsilon}\vec{\varepsilon}^T] \psi BA^T = \sigma_\varepsilon^2 ABA^T. \end{aligned}$$

Подставляя эти выражения в $R_{\hat{\theta}}$, получаем

$$R_{\hat{\theta}} = R_{\hat{\theta}} - 2\sigma_\varepsilon^2 BA^T CAB + \sigma_\varepsilon^2 BA^T \underbrace{CABA^T}_{=I} CAB = R_{\hat{\theta}} - \sigma_\varepsilon^2 BA^T CAB,$$

как и было записано ранее.

Этот метод часто применяется при решении геофизических задач. Например, в случае аппроксимации пространственной зависимости скорости распространения продольных волн, полученной по поверхностным наблюдениям (\vec{u}), данные акустического каротажа (полученные в скважинах) по определению скорости продольных волн (\vec{V}), могут быть использованы в качестве линейных ограничений.

6.5 Метод максимального правдоподобия (ММП)

Метод наименьших квадратов свободен относительно предположений относительно типа распределения вектора измерений \vec{u} .

В случае оценивания по *методу максимального правдоподобия*, плотность вероятностей измеренных величин должна быть определена.

Мы называем вероятность измеренных величин при заданных параметрах *правдоподобием* параметров при заданных измерениях.

Пусть случайный вектор \vec{u} геофизических наблюдений имеет плотность вероятностей $f(\vec{x}, \vec{\theta})$, зависящую от неизвестных (фиксированных) параметров $\vec{\theta}$, то *функция правдоподобия* $L(\vec{u}, \vec{\theta})$ определяется следующим образом

$$L(\vec{u}, \vec{\theta}) = f(\vec{x}, \vec{\theta}).$$

Оценка параметров $\vec{\theta}$, доставляющих максимум функции $L(\vec{u}, \vec{\theta})$ или $\ln L(\vec{u}, \vec{\theta})$ называется *оценкой максимального правдоподобия*

$$\hat{\vec{\theta}}_{MLM} = \arg \max \ln L(u_1, \dots, u_n; \vec{\theta}).$$

Для нахождения экстремума вычислим производные от $\ln L$ по параметрам θ_s

$$\frac{\partial \ln L(u_1, \dots, u_n; \vec{\theta})}{\partial \theta_s} = 0.$$

Такие уравнения называются *уравнениями максимального правдоподобия*.

Свойства оценок максимального правдоподобия.

1. $\hat{\vec{\theta}}_{MLM}$ – *состоятельная оценка*.
2. $\hat{\vec{\theta}}_{MLM}$ – *асимптотически несмещенная оценка* $M[\hat{\vec{\theta}}_n] \stackrel{n \rightarrow \infty}{\cong} \theta$.
3. $\hat{\vec{\theta}}_{MLM}$ – *асимптотически нормальная оценка* $\sqrt{n}(\hat{\vec{\theta}}_n - \vec{\theta}) \stackrel{n \rightarrow \infty}{\in} N(0, I^{(F)}(\theta))$, где $I^{(F)}(\theta)$ – информационная матрица Фишера.
4. $\hat{\vec{\theta}}_{MLM}$ – *асимптотически эффективная оценка*

$$\sigma_{\hat{\theta}_n}^2 \stackrel{n \rightarrow \infty}{\cong} [I^{(F)}(\theta_n)]^{-1}$$

и в случае вектора параметров

$$R(\hat{\vec{\theta}}) \stackrel{n \rightarrow \infty}{\cong} [I^{(F)}(\hat{\vec{\theta}}_n)]^{-1}.$$

Пример 1: В случае *произвольной аддитивной модели* геофизических данных с нормальным распределением случайной компоненты можно записать

$$\vec{u} = \vec{f}(\vec{\theta}) + \vec{\varepsilon}, \quad \vec{u} = (u_1, \dots, u_n), \quad \varepsilon \in N(0, R_\varepsilon).$$

Возьмем логарифм от функции правдоподобия

$$l(\vec{u}, \vec{\theta}) = -\frac{n}{2} \ln 2\pi - \frac{1}{2} \ln R_\varepsilon - \frac{1}{2} (\vec{u} - \vec{f}(\vec{\theta}))^T R_\varepsilon^{-1} (\vec{u} - \vec{f}(\vec{\theta})),$$

пренебрегая членами не зависящими от параметров получаем

$$l_1(\vec{u}, \vec{\theta}) = -\frac{1}{2} (\vec{u} - \vec{f}(\vec{\theta}))^T R_\varepsilon^{-1} (\vec{u} - \vec{f}(\vec{\theta})).$$

В частном случае, если

$$R_\varepsilon = \sigma_\varepsilon^2 I,$$

имеем

$$l_1(\vec{u}, \vec{\theta}) = -\frac{1}{2\sigma_\varepsilon^2} (\vec{u} - \vec{f}(\vec{\theta}))^T (\vec{u} - \vec{f}(\vec{\theta})).$$

Максимизация функции правдоподобия эквивалентна минимизации квадратичной формы

$$(\vec{u} - \vec{f}(\vec{\theta}))^T (\vec{u} - \vec{f}(\vec{\theta})),$$

т.е. в случае нормального распределения случайной компоненты ε метод максимального правдоподобия приводит к тому же результату, что и метод наименьших квадратов. Уравнения максимального правдоподобия запишутся в виде

$$(\vec{u} - \vec{f}(\vec{\theta}))^T R_\varepsilon^{-1} \frac{\partial \vec{f}}{\partial \theta_s} = 0, \quad s = 1, 2, \dots, S.$$

Пример 2: В случае линейной модели с нормальным распределением случайной компоненты

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon}, \quad \varepsilon \in N(0, R_\varepsilon)$$

функцию правдоподобия и уравнения правдоподобия можно записать в виде

$$l_1(\vec{u}, \vec{\theta}) = -\frac{1}{2} (\vec{u} - \psi \vec{\theta})^T R_\varepsilon^{-1} (\vec{u} - \psi \vec{\theta})$$

и

$$(\vec{u} - \psi \vec{\theta})^T R_\varepsilon^{-1} \psi = 0$$

соответственно. Оценка максимального правдоподобия в явном виде

$$\hat{\vec{\theta}}_{MLM} = (\psi^T R_\varepsilon^{-1} \psi)^{-1} \psi^T R_\varepsilon^{-1} \vec{u}.$$

При этом ковариационная матрица в асимптотике имеет вид

$$R_{\hat{\theta}} = \sigma^2 (\psi^T \psi)^{-1}.$$

В этом случае оценка максимального правдоподобия совпадает с оценкой МНК.

Пример 3: В случае *линейной модели* и лапласовского распределения случайной компоненты

$$\vec{u} = \psi\vec{\theta} + \vec{\varepsilon}, \quad \varepsilon \in L(0, \sigma_\varepsilon^2)$$

функцию правдоподобия можно записать в виде

$$l_1(\vec{u}, \vec{\theta}) = -\frac{2}{\sigma_\varepsilon^2} |\vec{u} - \psi\vec{\theta}|.$$

Очевидно, что максимизация функции правдоподобия приведет к методу наименьших модулей.

6.6 Метод наименьших модулей

Рассмотрим линейную модель наблюдений

$$\vec{u} = \psi\vec{\theta} + \vec{\varepsilon}$$

которая содержит выбросы. Робастная оценка может быть построена минимизацией по норме L_1 (или минимизацией модулей отклонений)

$$\lambda(\vec{\theta}) = \sum_{i=1}^n v_i |u_i - \sum_{s=1}^S \psi_{is} \theta_s|,$$

где v_i весовые коэффициенты. Оценку можно записать в виде

$$\hat{\vec{\theta}} = \arg \min_{\vec{\theta}} \sum_{i=1}^n v_i |u_i - \sum_{s=1}^S \psi_{is} \theta_s|.$$

Для нахождения оценки мы вводим функцию зависящую от векторов $\vec{\theta}$ и $\vec{\rho}$

$$\lambda_1(\vec{\theta}, \vec{\rho}) = \sum_{i=1}^n \frac{v_i^2 |u_i - \sum_{s=1}^S \psi_{is} \theta_s|^2}{v_i |u_i - \sum_{s=1}^S \psi_{is} \rho_s|}.$$

Отметим, что

$$\lambda_1(\vec{\theta}, \vec{\theta}) = \lambda(\vec{\theta}).$$

Вид функции $\lambda_1(\vec{\theta}, \vec{\rho})$ позволяет организовать итерационную процедуру

Пусть задан начальный вектор $\vec{\theta}^{(0)}$. Подставим заданную величину $\vec{\theta}^{(0)}$ в λ_1 вместо $\vec{\rho}$, теперь

$$\lambda_1(\vec{\theta}, \vec{\theta}^{(0)}) = \sum_{i=1}^n w_i^0 |u_i - \sum_{s=1}^S \psi_{is} \theta_s|^2,$$

где

$$w_i^0 = \frac{v_i}{|u_i - \sum_{s=1}^S \psi_{is} \theta_s^{(0)}|}$$

– элементы диагональной матрицы

$$W^{(0)} = \text{diag}(w_1^0, w_2^0, \dots, w_n^0).$$

Воспользуемся методом наименьших квадратов для нахождения первой итерации оценки $\vec{\theta}$

$$\begin{aligned}\hat{\vec{\theta}} &= (\psi^T W^{(0)} \psi)^{-1} \psi^T W^{(0)} \vec{u}, \\ \hat{\vec{\theta}}^{(1)} &= \hat{\vec{\theta}}.\end{aligned}$$

Используя полученную оценку $\hat{\vec{\theta}}^{(1)}$, найдем новые весовые коэффициенты $w_i^{(1)}$ для функции $\lambda_1(\vec{\theta}, \vec{\theta}^{(1)})$

$$\lambda_1(\vec{\theta}, \vec{\theta}^{(1)}) = \sum_{i=1}^n w_i^{(1)} (u_i - \sum_{s=1}^S \psi_{is} \theta_s)^2,$$

где

$$w_i^{(1)} = v_i / |u_i - \sum_{s=1}^S \psi_{is} \theta_s^{(1)}|.$$

Вторая итерация при вычислении оценки дается формулами

$$\begin{aligned}\hat{\vec{\theta}} &= (\psi^T W^{(1)} \psi)^{-1} \psi^T W^{(1)} \vec{u}, \\ \vec{\theta}^{(2)} &= \hat{\vec{\theta}}.\end{aligned}$$

После подстановки $\vec{\theta}^{(2)}$ вместо вектора $\vec{\rho}$ в функцию $\lambda_1(\vec{\theta}, \vec{\rho})$ найдем оценку $\vec{\theta}^{(3)}$.

Итерационный процесс следует продолжать до тех пор пока пороговые условия не будут удовлетворены

$$\frac{|\theta_s^{(k)} - \theta_s^{(k-1)}|}{|\theta_s^{(k)}|} < \delta \sim 10^{-2} \div 10^{-3}, \quad s = 1, \dots, S.$$

Основным элементом рассмотренной процедуры является взвешенный метод наименьших квадратов.

6.7 Робастный метод Хьюбера

Следующий робастный метод был предложен Хьюбером. Метод является комбинацией метода наименьших квадратов (для относительно малых случайных отклонений ε_i) и метода наименьших модулей для (больших величин ε_i). Рассмотрим линейную модель

$$u_i = \vec{\psi}_i^T \vec{\theta} + \varepsilon_i.$$

Оценка дается формулой

$$\hat{\vec{\theta}} = \arg \min_{\vec{\theta}} \sum_{i=1}^n H(u_i - \vec{\psi}_i^T \vec{\theta})$$

где

$$H(\varepsilon) = \begin{cases} \varepsilon^2/2 & |\varepsilon| < c, \\ c|\varepsilon| - c^2/2 & |\varepsilon| \geq c, \end{cases}$$

(см. рис. 6.1)

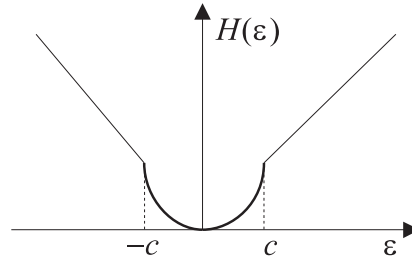


Рис. 6.1: Функция Хьюбера для расхождения между наблюдаемыми и модельными величинами.

6.8 Робастный метод Эндрюса

Рассмотрим робастный алгоритм предложенный Эндрюсом. Основным элементом этого алгоритма является взвешенный метод наименьших квадратов. Оценку получают минимизацией нелинейной функции

$$\hat{\vec{\theta}} = \arg \min_{\vec{\theta}} \sum_{i=1}^n H \left(\frac{u_i - \vec{\psi}_i^T \vec{\theta}}{\sigma} \right),$$

где

$$H(\varepsilon) = \begin{cases} 1 - \cos(\varepsilon/c) & |\varepsilon| < c\pi, \\ 0 & |\varepsilon| \geq c\pi, \end{cases}$$

$$\varepsilon_i = \frac{u_i - \vec{\psi}_i^T \vec{\theta}}{\sigma}$$

(see Figure 6.2). В качестве оценки среднеквадратического отклонения σ возьмем медиану модуля

$$\hat{\sigma} = \text{med} |u_i - \vec{\psi}_i^T \vec{\theta}|.$$

Решение задачи находят как решение системы уравнений

$$\sum_{i=1}^n h(\varepsilon_i) \psi_{is} = 0, \quad s = 1, \dots, S$$

где

$$h(\varepsilon_i) = H'(\varepsilon_i) = \begin{cases} \sin(\varepsilon_i/c)/c & |\varepsilon_i| < c\pi, \\ 0 & |\varepsilon_i| \geq c\pi. \end{cases}$$

Для решения нелинейной системы уравнений может быть использована следующая итерационная процедура:

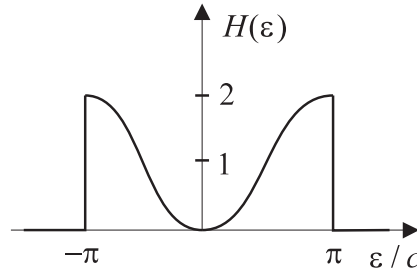


Рис. 6.2: Функция Эндрюса для расхождения между измеренными и модельными величинами.

1. Начальный вектор $\vec{\theta}_0$ может быть задан или получен с использованием МНК

$$\hat{\theta}^{(0)} = (\psi^T \psi)^{-1} \psi^T \vec{u}.$$

Рассмотрим алгоритм получения оценки $\hat{\theta}_0^{(k+1)}$ по известному значению $\hat{\theta}_0^{(k)}$, полученному на предыдущем шаге.

2. Вычислим отклонение

$$\hat{\varepsilon}_i^{(k)} = u_i - \vec{\psi}_i^T \hat{\theta}^{(k)}.$$

3. Найдем оценку для среднеквадратического отклонения σ

$$\hat{\sigma}^{(k)} = \text{med } \hat{\varepsilon}_i^{(k)}.$$

4. Вычислим весовые коэффициенты

$$w_i^{(k)} = \begin{cases} [\sin(\hat{\varepsilon}_i^{(k)}/c)/c][\hat{\varepsilon}_i^{(k)}]^{-1} & |\varepsilon_i|/c < \pi, \\ 0 & |\varepsilon_i|/c \geq \pi. \end{cases}$$

5. Решаем систему нелинейных уравнений

$$\sum_{i=1}^n w_i^k \psi_{is} \hat{\varepsilon}_i^{(k+1)} = 0$$

or

$$\sum_{i=1}^n w_i^k \psi_{is} (u_i - \vec{\psi}_i^T \vec{\theta}^{(k+1)}) = 0$$

и, окончательно, the $(k+1)$ -я итерация оценки дается выражением

$$\hat{\theta}^{(k+1)} = (\psi^T W^{(k)} \psi)^{-1} \psi^T W^{(k)} \vec{u}.$$

6. Проверяем пороговое условие

$$\frac{|\hat{\theta}_s^{(k+1)} - \hat{\theta}_s^{(k)}|}{|\hat{\theta}_s^{(k)}|} \leq \delta_1 \sim 10^{-2} \div 10^{-3}.$$

Если условие выполнено, то прекращаем итерационный процесс, иначе переходим к $(k+2)$ -й итерации.

6.9 Метод Байеса и статистическая регуляризация

Метод Байеса можно рассматривать как обобщение метода максимального правдоподобия. В методе Байеса априорная информация относительно искомых параметров вводится (формально) в оптимизационную процедуру. Концепция метода Байеса состоит в том, что плотность вероятностей $p(\vec{\theta}/\vec{u})$ (апостериорная плотность) для параметров $\vec{\theta}$ вычисляется с использованием наблюдений \vec{u} , плотности вероятностей $p(\vec{\theta})$ (априорная плотность) параметров $\vec{\theta}$ (которые предполагаются случайными), плотности $p(\vec{u}/\vec{\theta})$ наблюдений \vec{u} при заданных параметрах $\vec{\theta}$ (функция правдоподобия) и плотности $p(\vec{u})$ наблюдений \vec{u} :

$$p(\vec{\theta}/\vec{u}) = \frac{p(\vec{\theta})p(\vec{u}/\vec{\theta})}{p(\vec{u})}.$$

Имея апостериорную плотность можно определить неизвестные параметры $\vec{\theta}$.

Байесовское точечное оценивание определяется как математическое ожидание $\vec{\theta}$, когда весовой функцией является апостериорная плотность

$$\hat{\vec{\theta}}_B = \int \vec{\theta} p(\vec{\theta}/\vec{u}) d\vec{\theta}.$$

По аналогии со стандартным статистическим подходом, *обобщенную максимально правдоподобную оценку* $\hat{\vec{\theta}}_{MAP}$ вектора параметров $\vec{\theta}$ определяют как *моду* апостериорной плотности $p(\vec{\theta}/\vec{u})$, т.е. это величина $\vec{\theta}$ при которой апостериорная плотность $p(\vec{\theta}/\vec{u})$ достигает максимального значения

$$\hat{\vec{\theta}}_{MAP} = \arg \max_{\vec{\theta}} p(\vec{\theta}/\vec{u}).$$

Учитывая, что логарифм является монотонной функцией:

$$\hat{\vec{\theta}}_{MAP} = \arg \max_{\vec{\theta}} \ln p(\vec{\theta}/\vec{u}).$$

Если максимум лежит в допустимом интервале для параметров $\vec{\theta}$ и функция $\ln p(\vec{\theta}/\vec{u})$ дифференцируема, оценку по максимуму апостериорной вероятности можно находить решая систему уравнений

$$\frac{\partial \ln p(\vec{\theta}/\vec{u})}{\partial \theta_s} = 0, \quad s = 1, \dots, S$$

или, принимая во внимание формулу Байеса для $\ln p(\vec{\theta}/\vec{u})$

$$\ln p(\vec{\theta}/\vec{u}) = \ln p(\vec{\theta}) + \ln p(\vec{u}/\vec{\theta}) - \ln p(\vec{u})$$

можно записать

$$\frac{\partial \ln p(\vec{\theta})}{\partial \theta_s} + \frac{\partial \ln p(\vec{u}/\vec{\theta})}{\partial \theta_s} = 0, \quad s = 1, 2, \dots, S.$$

Первое слагаемое связано с априорной вероятностью, а второе – с наблюдениями.

Если функция $\ln p(\vec{\theta})$ слабо изменяется в области допустимых изменений параметров $\vec{\theta}$, то мы можем пренебречь первым слагаемым. Например, если априорная плотность $p(\vec{\theta})$ является постоянной величиной $p(\vec{\theta}) = \text{const}$, то оценка по максимуму апостериорной вероятности совпадает с максимально правдоподобной оценкой.

В общем случае, введение априорной вероятности (не равной константе) приводит к устойчивости оценок, полученных по максимуму апостериорной вероятности. Использование априорной информации лежит в основе метода *статистической регуляризации*.

Пример: Найдём оценку по максимуму апостериорной вероятности в случае линейной модели

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon}$$

и нормального распределения $\vec{\varepsilon}$ и $\vec{\theta}$

$$\vec{\varepsilon} \in N(0, R_\varepsilon), \quad \vec{\theta} \in N(\langle \vec{\theta} \rangle, R_\theta).$$

В этом случае можно записать

$$\begin{aligned} \ln p(\vec{\theta}/\vec{u}) &= -\frac{1}{2}(\vec{u} - \psi \vec{\theta})^T R_\varepsilon^{-1}(\vec{u} - \psi \vec{\theta}) - \frac{1}{2}(\vec{\theta} - \langle \vec{\theta} \rangle)^T R_\theta^{-1} \times \\ &\times (\vec{\theta} - \langle \vec{\theta} \rangle) - \frac{n+S}{2} \ln(2\pi) - \frac{1}{2}(\ln |R_\theta| + \ln |R_\varepsilon|) - \ln p(\vec{u}). \end{aligned}$$

Система уравнений для получения оценки по методу максимума апостериорной вероятности имеет вид

$$\frac{\partial \ln p(\vec{\theta}/\vec{u})}{\partial \theta_s} = 0, \quad s = 1, 2, \dots, S$$

и подставляя выражение для $\ln p(\vec{\theta}/\vec{u})$, получаем

$$(\psi^T R_\varepsilon^{-1} \psi + R_\theta^{-1}) \vec{\theta} = \psi^T R_\varepsilon^{-1} \vec{u} + R_\theta^{-1} \langle \vec{\theta} \rangle.$$

Оценка по максимуму апостериорной вероятности записывается в виде

$$\vec{\theta}_{MAP}^{\sim} = (\psi^T R_\varepsilon^{-1} \psi + R_\theta^{-1})^{-1} (\psi^T R_\varepsilon^{-1} \vec{u} + R_\theta^{-1} \langle \vec{\theta} \rangle),$$

при этом ковариационная матрица оценки определяется формулой

$$R_{\vec{\theta}_{MAP}^{\sim}} = (\psi^T R_\varepsilon^{-1} \psi + R_\theta^{-1})^{-1}.$$

Используя равенства

$$\begin{aligned}(\psi^T R_\varepsilon^{-1} \psi + R_\theta^{-1})^{-1} &= R_\theta - R_\theta \psi^T (\psi R_\theta \psi^T + R_\varepsilon)^{-1} \psi R_\theta, \\ (\psi^T R_\varepsilon^{-1} \psi + R_\theta^{-1})^{-1} \psi^T R_\varepsilon^{-1} &= R_\theta \psi^T (\psi R_\theta \psi^T + R_\varepsilon)^{-1},\end{aligned}$$

запишем иначе выражение для оценки

$$\tilde{\theta}_{MAP} = \langle \vec{\theta} \rangle + R_\theta \psi^T (\psi R_\theta \psi^T + R_\varepsilon)^{-1} (\vec{u} - \psi \langle \vec{\theta} \rangle)$$

и ее ковариационной матрицы

$$R_{\tilde{\theta}_{MAP}} = R_\theta - R_\theta \psi^T (\psi R_\theta \psi^T + R_\varepsilon^{-1})^{-1} \psi R_\theta.$$

В частном случае некоррелированной случайной компоненты $\vec{\varepsilon}$

$$R_\varepsilon = \sigma_\varepsilon^2 I_n,$$

и некоррелированных компонент вектора $\vec{\theta}$

$$R_\theta = \sigma_\theta^2 I_S$$

рассмотренные ранее выражения для оценок запишутся следующим образом

$$\tilde{\theta}_{MAP} = (\psi^T \psi + \alpha I_S)^{-1} (\psi^T \vec{u} + \alpha \langle \vec{\theta} \rangle)$$

и

$$\tilde{\theta}_{MAP} = \langle \vec{\theta} \rangle + \psi^T (\psi \psi^T + \alpha I_n)^{-1} (\vec{u} - \psi \langle \vec{\theta} \rangle),$$

где $\alpha = \sigma_\varepsilon^2 / \sigma_\theta^2$. Соответствующие ковариационные матрицы:

$$R_{\tilde{\theta}_{MAP}} = \sigma_\varepsilon^2 (\psi^T \psi + \alpha I_S)^{-1}$$

и

$$R_{\tilde{\theta}_{MAP}} = \sigma_\theta^2 (I_S - \psi^T (\psi \psi^T + \alpha I_n)^{-1} \psi).$$

Заметим, что параметр α можно интерпретировать как параметр *регуляризации по Тихонову*. В случае статистической регуляризации параметр α имеет очень прозрачный смысл – это энергетическое соотношение “шум/сигнал”. При стремлении α к нулю, оценка по максимуму апостериорной вероятности $\tilde{\theta}_{MAP}$ стремится к оценке максимального правдоподобия $\tilde{\theta}_{MLS}$.

6.9.1 Рекурсивный алгоритм для метода максимума апостериорной вероятности

В случае представления оценки MAP и ее ковариационной матрицы соответственно в виде

$$\tilde{\theta}_{MAP} = \langle \vec{\theta} \rangle + R_\theta \psi^T (\psi R_\theta \psi^T + R_\varepsilon)^{-1} (\vec{u} - \psi \langle \vec{\theta} \rangle)$$

и

$$R_{\tilde{\theta}_{MAP}} = R_{\theta} - R_{\theta}\psi^T(\psi R_{\theta}\psi^T + R_{\varepsilon})^{-1}\psi R_{\theta}$$

можно построить рекурсивный алгоритм для нахождения как оценки, так и ковариационной матрицы.

В качестве нулевого приближения для вектора параметров возьмем математическое ожидание от параметров по априорной плотности вероятностей $\langle \vec{\theta} \rangle$, а в качестве нулевого приближения для ковариационной матрицы – априорную ковариационную матрицу R_{θ}

$$\vec{\theta}^{(0)} = \langle \vec{\theta} \rangle, \quad R_{\theta}^{(0)} = R_{\theta}.$$

Будем предполагать независимость измерений:

$$R_{\varepsilon} = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2).$$

Первое приближение для искомой оценки получаем по формуле

$$\vec{\theta}^{(1)} = \vec{\theta}^{(0)} + \gamma^{(1)} \Delta \vec{\theta}^{(1)},$$

где

$$\begin{aligned} \Delta \vec{\theta}^{(1)} &= R_{\theta}^{(0)} \vec{\psi}_1^T, \quad \gamma^{(1)} = N_1 / \delta_1^2, \\ N_1 &= u_1 - \vec{\psi}_1^T \vec{\theta}^{(0)}, \quad \delta_1^2 = \sigma_{\varepsilon_1}^2 + \vec{\psi}_1^T \Delta \vec{\theta}^{(1)}, \quad \vec{\psi}_1 = \|\psi_{11}, \psi_{12}, \dots, \psi_{1S}\|. \end{aligned}$$

Первое приближение для ковариационной матрицы дается выражением

$$R_{\theta}^{(1)} = R_{\theta}^{(0)} - \frac{\Delta \vec{\theta}^{(1)} \Delta \vec{\theta}^{(1)T}}{\delta_1^2}.$$

Пусть имеется оценка и ее ковариационная матрица для $(i-1)$ -о шага

$$\vec{\theta}^{(i-1)}, \quad R_{\theta}^{(i-1)},$$

тогда оценку и ее ковариационную матрицу на i -м шаге можно записать в виде

$$\vec{\theta}^{(i)} = \vec{\theta}^{(i-1)} + \gamma^i \Delta \vec{\theta}^{(i)},$$

где

$$\begin{aligned} \Delta \vec{\theta}^{(i)} &= R_{\theta}^{(i-1)} \vec{\psi}_i^T, \quad \gamma^{(i)} = N_i / \delta_i^2, \\ N_i &= u_i - \vec{\psi}_i^T \vec{\theta}^{(i-1)}, \quad \delta_i^2 = \sigma_{\varepsilon_i}^2 + \vec{\psi}_i^T \Delta \vec{\theta}^{(i)} \end{aligned}$$

и

$$R_{\theta}^{(i)} = R_{\theta}^{(i-1)} - \Delta \vec{\theta}^{(i)} \Delta \vec{\theta}^{(i)T} / \delta_i^2.$$

Следует отметить, что такая процедура является стабильной и может быть легко реализована для любого объема исходных данных. Эта процедура может быть использована для динамической или калмановской фильтрации.

6.9.2 Фильтр Калмана

В процессе распространения сейсмической волны происходит изменение ее формы. Мы можем учесть это изменение используя следующую модель

$$\begin{aligned} u_i &= \vec{\varphi}_i \vec{\theta}_i + \varepsilon_i, \quad \varepsilon_i \in N(0, \sigma_{\varepsilon_i}^2), \\ \vec{\theta}_i &= F_{ii-1} \vec{\theta}_{i-1} + \vec{G} \delta_{i-1}, \quad \delta_i \in N(0, \sigma_{\delta_i}^2), \end{aligned}$$

где векторы амплитуд записываются в виде

$$\vec{\theta}_i = \begin{bmatrix} \theta_i^{(1)} \\ \vdots \\ \theta_i^{(M)} \end{bmatrix}, \quad \vec{\theta}_{i-1} = \begin{bmatrix} \theta_{i-1}^{(1)} \\ \vdots \\ \theta_{i-1}^{(M)} \end{bmatrix}, \quad \vec{G} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Компоненты векторов θ_i and θ_{i-1} связаны соотношениями

$$\begin{bmatrix} \theta_i^{(2)} = \theta_{i-1}^{(1)} \\ \dots \\ \theta_i^{(M)} = \theta_{i-1}^{(M-1)} \end{bmatrix}.$$

Вектор $\vec{\varphi}_i = [\varphi_i^{(1)}, \dots, \varphi_i^{(M)}]$ – вектор формы сигнала (см. рис. 6.3). Матрица перехода от вектора $\vec{\theta}_{i-1}$ к вектору $\vec{\theta}_i$ дается выражением

$$F_{ii-1} = \begin{bmatrix} r_{i-1}(1) & r_{i-1}(2) & \dots & r_{i-1}(M-1) & r_{i-1}(M) \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

где $r_{i-1}(\mu)$ – элементы априорной ковариационной функции.

Будем получать оценку $\vec{\theta}_i$ максимизируя апостериорную плотность

$$\ln p(\vec{\theta}_i / \vec{u})$$

или, учитывая нормальное распределение случайных компонент ε_i и δ_i , можно записать следующее

$$l(\vec{\theta}_i) = \frac{-1}{2\sigma_{\varepsilon_i}^2} (u_i - \vec{\varphi}_i \vec{\theta}_i)^2 - \frac{1}{2} (\vec{\theta}_i - \hat{\vec{\theta}}_i')^T R_{\theta_i'}^{-1} (\vec{\theta}_i - \hat{\vec{\theta}}_i'),$$

где $\hat{\vec{\theta}}_i' = F_{ii-1} \vec{\theta}_i$, при этом ковариационная матрица:

$$R_{\theta_i'}^{-1} = \langle \vec{\theta}_i \vec{\theta}_i^T \rangle = F_{ii-1} R_{\theta_{i-1}} F_{ii-1}^T + \vec{G} \sigma_{\delta_{i-1}}^2 \vec{G}^T.$$

Оценка может быть получена как результат максимизации функции $l(\vec{\theta}_i)$

$$\tilde{\vec{\theta}}_i = \left(\frac{\vec{\varphi}_i^T \vec{\varphi}_i}{\sigma_{\varepsilon_i}^2} + R_{\theta_i'}^{-1} \right)^{-1} \left(\frac{\vec{\varphi}_i^T u_i}{\sigma_{\varepsilon_i}^2} + R_{\theta_i'}^{-1} \hat{\vec{\theta}}_i' \right)$$

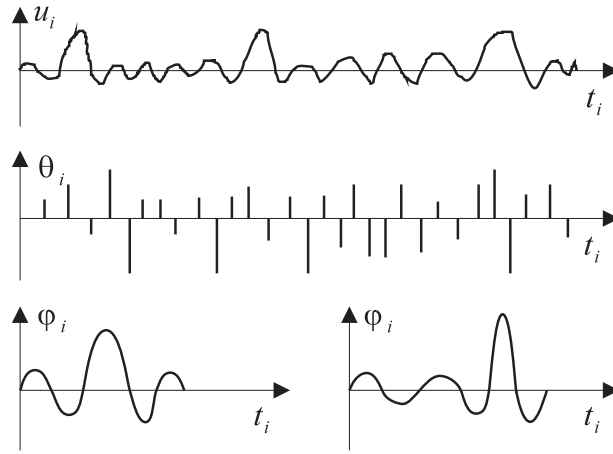


Рис. 6.3: Иллюстрация процедуры калмановской фильтрации.

или, используя равенства рассмотренные ранее, получим рекурсивные формулы для оценки

$$\tilde{\theta}_i = \hat{\theta}_i' + R_{\theta_i'} \bar{\varphi}_i^T (\bar{\varphi}_i R_{\theta_i'} \bar{\varphi}_i^T + \sigma_{\varepsilon i}^2)^{-1} (u_i - \bar{\varphi}_i \hat{\theta}_i')$$

и ее ковариационной матрицы

$$R_{\theta_i} = R_{\theta_i'} - R_{\theta_i'} \bar{\varphi}_i^T (\bar{\varphi}_i R_{\theta_i'} \bar{\varphi}_i^T + \sigma_{\varepsilon i}^2)^{-1} \bar{\varphi}_i R_{\theta_i'}$$

Начальные значения для оценки и ковариационной матрицы можно, например, взять в виде

$$\vec{\theta}_0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \quad R_{\theta_0} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

6.10 Сингулярный анализ и метод наименьших квадратов

Рассмотрим линейную модель

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon}, \quad \varepsilon \in N(0, \sigma_\varepsilon^2).$$

При практическом применении метода наименьших квадратов не удается получить сколько-нибудь удовлетворительный результат в случае плохой обусловленности матрицы $\psi^T \psi$, которая подвергается обращению. В качестве меры обусловленности обычно используют *число обусловленности*

$$\beta = G/g,$$

где

$$G = \max_{\vec{\theta}} \frac{\|\psi^T \psi \vec{\theta}\|_1}{\|\vec{\theta}\|_1}, \quad g = \min_{\vec{\theta}} \frac{\|\psi^T \psi \vec{\theta}\|_1}{\|\vec{\theta}\|_1}, \quad \|\vec{\theta}\|_1 = \sum_{s=1}^S |\theta_s|.$$

Если матрица $\psi^T\psi$ – сингулярная, то величина g равна нулю.

Свойства чисел обусловленности.

1. Число обусловленности β всегда больше или равно единице

$$\beta \geq 1 \quad \text{as} \quad G \geq g.$$

2. Если матрица $\psi^T\psi$ – диагональная, как в случае ортогональных полиномов, то

$$\beta = \frac{\max |(\psi^T\psi)_{ii}|}{\min |(\psi^T\psi)_{ii}|}.$$

3. Если существует обратная матрица, то

$$\beta = \|\psi^T\psi\| \|(\psi^T\psi)^{-1}\|.$$

4. Если матрица $\psi^T\psi$ – сингулярная, то

$$\beta = \infty.$$

Большое значение числа обусловленности свидетельствует о “почти линейной зависимости” строк матрицы.

Произвольная вещественная матрица ψ может быть представлена своим сингулярным разложением

$$\psi = Q\Sigma P^T,$$

где $Q_{m \times n}$ и $P_{S \times S}$ – ортогональные матрицы, т.е.

$$QQ^T = I_{n \times n}, \quad PP^T = I_{S \times S}.$$

Матрица $\Sigma_{n \times S}$ – диагональная с элементами

$$\Sigma_{ij} = \begin{cases} \sigma_i & i = j, \quad i = 1, \dots, n, \\ 0 & i \neq j, \quad j = 1, \dots, S, \end{cases} \quad \Sigma = \underbrace{\begin{bmatrix} \sigma_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \sigma_S \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{bmatrix}}_S \Bigg\} n. \quad (6.1)$$

Матрица Q составлена из ортонормированных собственных векторов матрицы $\psi\psi^T$, и матрица P составлена из ортонормированных векторов матрицы $\psi^T\psi$. Диагональные элементы матрицы Σ неотрицательные значения квадратных корней из собственных значений матрицы $\psi^T\psi$, они называются *сингулярными величинами*. Будем считать выполненными следующие неравенства

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_S \geq 0.$$

Для простоты рассмотрим некоррелированную случайную компоненту ε . В этом случае оценка по МНК имеет вид

$$\hat{\vec{\theta}} = \arg \min \lambda(\vec{\theta}),$$

где

$$\begin{aligned} \lambda(\vec{\theta}) &= (\vec{u} - \psi\vec{\theta})^T (\vec{u} - \psi\vec{\theta}) = (\vec{u} - Q\Sigma P^T \vec{\theta})^T Q Q^T \times \\ &\times (\vec{u} - Q\Sigma P^T \vec{\theta}) = (\vec{y} - \Sigma\vec{b})^2, \end{aligned}$$

где $\vec{y} = Q^T \vec{u}$ и $\vec{b} = P^T \vec{\theta}$.

Минимизация функции λ дает оценку

$$\hat{\vec{b}} = \Sigma^+ \vec{y}$$

и

$$\hat{\vec{\theta}} = P \Sigma^+ Q^T \vec{u}.$$

Матрицу Σ^+ называют псевдообратной матрицей: as

$$\Sigma^+ = \underbrace{\begin{bmatrix} \overbrace{\sigma_1^{-1} \dots 0}^S & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \sigma_S^{-1} & 0 \end{bmatrix}}_n \underbrace{\begin{bmatrix} 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{bmatrix}}_{n-S}.$$

Свойства псевдообратной матрицы.

1. $\Sigma \Sigma^+ \Sigma = \Sigma$.
2. $\Sigma^+ \Sigma \Sigma^+ = \Sigma^+$.
3. $\Sigma \Sigma^+$ – симметричная матрица.
4. $\Sigma^+ \Sigma$ – симметричная матрица.

Можно показать, что матрица Σ^+ всегда существует и единственна. Если диагональные элементы $\sigma_{ii} \neq 0$, то оценка легко находится

$$b_i = \frac{y_i}{\sigma_{ii}}, \quad i = 1, 2, \dots, i_0.$$

При малых величинах σ_{ii} такой путь может не привести к сколько-нибудь удовлетворительному результату. Следует выбрать пороговую величину α , отталкиваясь от априорной точности исходных данных. Затем из сравнения значений σ_{ii} и α сделать следующий выбор.

Если $\sigma_{ii} \geq \alpha$, то $b_i = y_i / \sigma_{ii}$ для $i = 1, 2, \dots, i_0$.

Если $\sigma_{ii} < \alpha$, то $b_i = 0$ для $i = i_0 + 1, i_0 + 2, \dots, S$.

Такая процедура носит название *сингулярного анализа*.

Пример: Применим сингулярный анализ для оценивания в случае случайного вектора параметров $\vec{\theta}$ с

$$\langle \vec{\theta} \rangle = 0 \quad \text{and} \quad R_{\theta} = \sigma_{\theta}^2 I.$$

Полагаем случайный вектор $\vec{\varepsilon}$ некоррелированным

$$R_{\varepsilon} = \sigma_{\varepsilon}^2 I.$$

Полагаем также некоррелированными случайные векторы $\vec{\varepsilon}$ и $\vec{\theta}$.

Воспользовавшись сингулярным разложением матрицы ψ запишем модель в виде

$$\vec{y} = \Sigma \vec{b} + \vec{\varepsilon},$$

где $\vec{e} = Q^T \vec{\varepsilon}$. Решение можно записать в виде линейной формы

$$\hat{\vec{b}} = \Gamma \vec{y}, \tag{6.2}$$

где матрица Γ – диагональная матрица $[S \times n]$

$$\Gamma = \left. \begin{array}{c} \overbrace{\left[\begin{array}{cccccc} \gamma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \gamma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \gamma_S & 0 & \dots & 0 \end{array} \right]}^n \\ \underbrace{\hspace{10em}}_S \quad \underbrace{\hspace{10em}}_{n-S} \end{array} \right\} \begin{array}{l} S \\ \gamma_1 \\ \gamma_1 \\ \vdots \\ \gamma_s \end{array} - ?.$$

Для нахождения оценки $\hat{\vec{b}}$ достаточно найти элементы $\gamma_1, \dots, \gamma_S$.

Воспользуемся критерием наименьших квадратов. Ошибка оценки и ее ковариационная матрица даются формулами

$$\vec{\xi} = \hat{\vec{b}} - \vec{b} = (\Gamma \Sigma - I) \vec{b} + \Gamma \vec{\varepsilon},$$

и

$$\begin{aligned} \langle \vec{\xi} \vec{\xi}^T \rangle &= (\Gamma \Sigma - I) \langle \vec{b} \vec{b}^T \rangle (\Gamma \Sigma - I)^T + \\ &+ \Gamma \langle \vec{\varepsilon} \vec{\varepsilon}^T \rangle \Gamma^T = (\Gamma \Sigma - I) R_b (\Gamma \Sigma - I)^T + \Gamma R_{\varepsilon} \Gamma^T. \end{aligned}$$

Сумма квадратов ошибок:

$$\sigma_{\xi}^2 = \text{Sp}(\langle \vec{\xi} \vec{\xi}^T \rangle) = \sum_{j=1}^S [(\gamma_j \sigma_j - 1)^2 \sigma_{b_j}^2 + \gamma_j^2 \sigma_{\varepsilon_j}^2].$$

Принимая во внимание, что

$$R_b = P^T \langle \vec{\theta} \vec{\theta}^T \rangle P = \sigma_{\theta}^2 I$$

и

$$R_e = Q^T \langle \vec{\xi} \vec{\xi}^T \rangle Q = \sigma_\varepsilon^2 I,$$

можно записать

$$\sigma_\xi^2 = \sum_{j=1}^S [(\gamma_j \sigma_j - 1)^2 \sigma_\theta^2 + \gamma_j^2 \sigma_\varepsilon^2].$$

Дифференцируя σ_ξ^2 по γ_j

$$\begin{aligned} \frac{\partial \sigma_\xi^2}{\partial \gamma_j} &= 0, \quad j = 1, \dots, S, \\ 2\sigma_j(\gamma_j \sigma_j - 1)\sigma_\theta^2 + 2\gamma_j \sigma_\varepsilon^2 &= 0 \end{aligned}$$

получим оценку $\hat{\gamma}_j$

$$\hat{\gamma}_j = \frac{\sigma_j}{(\sigma_j^2 + \sigma_\varepsilon^2 / \sigma_\theta^2)},$$

доставляющую регуляризованную оценку \hat{b} . Такая регуляризация называется *статистической регуляризацией*. В частном случае $\sigma_\theta \rightarrow \infty$, получаем оценку по *методу наименьших квадратов*

$$\hat{\gamma}_j = \frac{1}{\sigma_j}.$$

В случае сингулярного разложения сумму квадратов ошибок можно представить в виде

$$\lambda(i_0) = \langle \|\xi\|^2 \rangle = \sigma_\theta^2 (S - i_0) + \sigma_\varepsilon^2 \sum_{j=1}^{i_0} 1/\sigma_j^2.$$

Максимизируя $\lambda(i_0)$ по отношению i_0 получаем

$$\sigma_{i_0} \geq \sigma_\varepsilon^2 / \sigma_\theta^2 = \alpha \quad \text{and} \quad \sigma_{i_0+1} < \sigma_\varepsilon^2 / \sigma_\theta^2 = \alpha,$$

$$\gamma_j = \begin{cases} 1/\sigma_j & j = 1, 2, \dots, i_0 \\ 0 & j = i_0 + 1, \dots, S \end{cases} \quad \begin{matrix} \sigma_{i_0} \geq \alpha, \\ \sigma_{i_0+1} < \alpha \end{matrix}$$

(см. рис. 6.4). Отметим, что сингулярный анализ является одним из подходов для регуляризации.

6.10.1 Матрица разрешения

Рассмотрим модель

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon}.$$

Оценка может быть записана в виде

$$\hat{\vec{\theta}} = L\vec{u}.$$

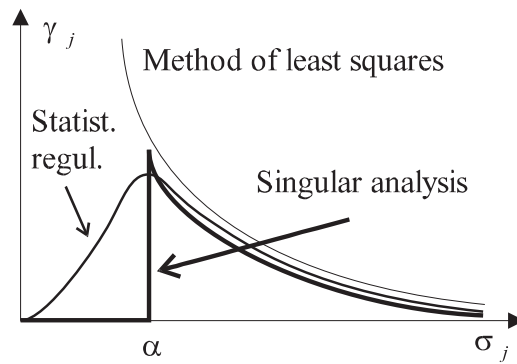


Рис. 6.4: Зависимость элементов матрицы Γ (6.2) от собственных чисел σ_j (6.1) для случаев метода наименьших квадратов, статистической регуляризации и сингулярного анализа.

Введем матрицу разрешения

$$R = L\psi \Rightarrow I_S.$$

В этом случае

$$\hat{\theta} \stackrel{\varepsilon \rightarrow 0}{\Rightarrow} R\vec{\theta}.$$

В случае оценивания по МНК оператор L записывается в виде

$$L = (\psi^T \psi)^{-1} \psi^T, \quad R = L\psi = I_S.$$

Рассмотрим сингулярное разложение

$$\psi = Q\Sigma P^T.$$

Матрица разрешения может быть записана:

$$R = LQ\Sigma P^T = CP^T.$$

Будем искать \vec{c}_j минимизируя квадратичную форму

$$\lambda(\vec{c}_j) = (P\vec{c}_j - \vec{\delta}_j)^T (P\vec{c}_j - \vec{\delta}_j),$$

где

$$\vec{\delta}_j = \|\delta_{ij}\|_{i,j=1}^S, \quad \delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j, \end{cases}$$

$$\vec{c}_j = (P^T P)^{-1} P^T \delta_j = P^T \vec{\delta}_j,$$

$$C = P, \quad R = PP^T.$$

6.11 Метод Бейкуса–Гильберта

Пусть $\theta(t)$ – неизвестная параметрическая функция, и мы рассматриваем конечное число измерений, u_1, u_2, \dots, u_n , которые являются линейными функционалами θ с ядром $\varphi_i(t)$:

$$u_i = \int_0^T \varphi_i(t)\theta(t)dt + \varepsilon_i, \quad \varepsilon_i \in N(0, R_\varepsilon), \quad i = 1, 2, \dots, n,$$

где ε_i – нормально распределенные случайные величины. Задача состоит в нахождении (восстановлении) функции $\theta(t)$.

Истинную (неизвестную) функцию обозначим как $\theta(t_0)$:

$$\langle \theta(t_0) \rangle = \int_0^T C(t_0, t)\theta(t)dt,$$

где функция $C(t_0, t)$ должна быть близка к δ -функции

$$C(t_0, t) = \delta(t - t_0) \Rightarrow \langle \theta(t_0) \rangle = \theta(t_0).$$

Предполагаем, что $C(t_0, t)$ – нормированная функция

$$\int_0^T C(t_0, t)dt = 1.$$

Пусть $C(t_0, t)$ – линейная комбинация функций $\varphi_i(t)$

$$C(t_0, t) = \sum_{i=1}^n \varphi_i(t)\rho_i(t_0),$$

где $\rho_i(t_0)$ – неизвестные коэффициенты. Подставив это выражение в выражение для оценки $\langle \theta(t_0) \rangle$, получим

$$\langle \theta(t_0) \rangle = \sum_{i=1}^n \rho_i(t_0)f_i,$$

где

$$f_i = \int_0^T \varphi_i(t)\theta(t)dt.$$

Модель можем записать в виде

$$u_i = f_i + \varepsilon_i.$$

Будем искать оценку как линейную комбинацию результатов измерений

$$\langle \hat{\theta}(t_0) \rangle = \sum_{i=1}^n \rho_i(t_0) u_i.$$

Теперь задача состоит в нахождении “наилучших” значений величин $\rho_i(t_0)$. Введем понятие *разрешающей длины*

$$s(t_0) = 12 \int_{t_0-L/2}^{t_0+L/2} (t-t_0)^2 C^2(t_0, t) dt,$$

$$C(t_0, t) = \begin{cases} 1/L & \text{for } t \in [t_0 - L/2 \leq t \leq t_0 + L/2] \\ 0 & \text{for } t \notin [t_0 - L/2 \leq t \leq t_0 + L/2]. \end{cases}$$

Функция $C(t_0, t)$ представлена на рис. 6.5. В этом случае получим

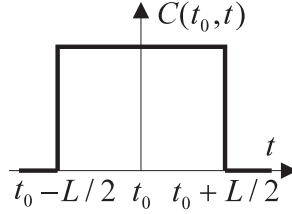


Рис. 6.5: Графическое представление функции $C(t_0, t)$.

$$s(t_0) = (12/3L^2)((L/2)^3 - (-L/2)^3) = L.$$

Величину $s(t_0)$ можно представить как квадратичную форму

$$s(t_0) = \vec{\rho}^T(t_0) S(t_0) \vec{\rho}(t_0),$$

где вектор $\vec{\rho}(t_0)$ и матрица $S(t_0) = \|S_{\mu\nu}(t_0)\|_{\mu,\nu}^n$ могут быть представить в виде

$$\vec{\rho}(t_0) = [\rho_1(t_0), \rho_2(t_0), \dots, \rho_n(t_0)],$$

$$S_{\mu\nu}(t_0) = 12 \int_{t_0-L/2}^{t_0+L/2} \varphi_\mu(t) \varphi_\nu(t) (t-t_0)^2 dt.$$

Ошибка восстановления имеет вид

$$\begin{aligned} \sigma_\rho^2(t_0) &= \langle (\langle \theta(t_0) \rangle - \langle \hat{\theta}(t_0) \rangle)^2 \rangle = \langle (\sum_{i=1}^n \rho_i(t_0) (u_i - f_i))^2 \rangle = \\ &= \vec{\rho}^T(t_0) R_\varepsilon \vec{\rho}(t_0). \end{aligned}$$

Окончательно, предлагается минимизировать сумму взвешенных квадратичных (относительно $\vec{\rho}(t_0)$) форм $s(t_0)$ и $\sigma_\rho^2(t_0)$

$$g(\vec{\rho}(t_0)) = s(t_0) \cos \theta + \sigma^2(t_0) \beta \sin \theta = \vec{\rho}^T(t_0) G \vec{\rho}(t_0), \quad 0 \leq \theta \leq \pi/2.$$

где $G = S \cos \theta + \beta R_\varepsilon \sin \theta$. Будем искать оценку $\vec{\rho}$

$$\hat{\vec{\rho}} = \arg \min g(\vec{\rho}(t_0))$$

при условии

$$\vec{B}^T \vec{\rho}(t_0) = 1, \quad \vec{B} = \|B_\nu\|_{\nu=1}^n, \quad B_\nu = \int_0^T \varphi_\nu(t) dt.$$

Это задача на условный экстремум

$$\begin{aligned} \Phi(\vec{\rho}) &= \vec{\rho}^T G \vec{\rho} - 2\lambda(\vec{\rho}^T \vec{B} - 1), \\ \frac{\partial \Phi}{\partial \rho_s} &= 0, \quad s = 1, 2, \dots, S. \end{aligned}$$

В результате получим систему линейных уравнений

$$G \vec{\rho} - \lambda \vec{B} = 0$$

и запишем оценку $\vec{\rho}$ в виде

$$\hat{\vec{\rho}} = \lambda G^{-1} \vec{B}.$$

Умножая последнее равенство на \vec{B}^T , получаем

$$\vec{B}^T \hat{\vec{\rho}} = \lambda \vec{B}^T G^{-1} \vec{B},$$

используя условие

$$\vec{B}^T \hat{\vec{\rho}} = 1,$$

записываем

$$\lambda = 1/(\vec{B}^T G^{-1} \vec{B}).$$

Окончательно, оценка $\vec{\rho}$ имеет вид

$$\hat{\vec{\rho}} = G^{-1} \vec{B} / (\vec{B}^T G^{-1} \vec{B}).$$

Глава 7

Решение обратных геофизических задач: нелинейные модели

7.1 Метод Ньютона Лекама

Рассмотрим модель измерения

$$\vec{u} = \vec{f}(\vec{\theta}) + \vec{\varepsilon}$$

$\varepsilon \in N(0, R_\varepsilon)$, $f(\theta)$ – нелинейная функция вектора параметров $\vec{\theta}$. Применение метода максимального правдоподобия приводит к минимизации квадратичной формы

$$l_1(\vec{u}, \vec{\theta}) = \frac{1}{2}(\vec{u} - \vec{f}(\vec{\theta}))^T R_\varepsilon^{-1}(\vec{u} - \vec{f}(\vec{\theta})).$$

Зададим начальный вектор $\vec{\theta}^{(0)}$ для вектора искомых параметров. После разложения функции $l_1(\vec{\theta})$ в ряд Тейлора в окрестности $\vec{\theta}^{(0)}$, получаем квадратичную форму относительно искомой величины $\vec{\theta}$

$$l_1(\vec{\theta}) \approx l_1(\vec{\theta}^{(0)}) + \Delta\vec{\theta}^T \vec{d} - (1/2)\Delta\vec{\theta}^T C \Delta\vec{\theta},$$

где

$$\Delta\vec{\theta} = \vec{\theta} - \vec{\theta}_0, \quad d_s = \left. \frac{\partial l_1(\vec{\theta})}{\partial \theta_s} \right|_{\vec{\theta}=\vec{\theta}^{(0)}}, \quad c_{ss'} = -\left. \frac{\partial^2 l_1(\vec{\theta})}{\partial \theta_s \partial \theta_{s'}} \right|_{\vec{\theta}=\vec{\theta}^{(0)}}.$$

Первый шаг итерационной процедуры состоит в решении уравнения

$$\frac{\partial l_1(\vec{\theta})}{\partial \theta_s} = 0 \Rightarrow C \Delta\vec{\theta} = \vec{d}, \quad s = 1, 2, \dots, S.$$

Лекам предложил использовать математическое ожидание элементов $c_{ss'}$ матрицы C

$$\tilde{c}_{ss'}^{(0)} = \langle c_{ss'}^{(0)} \rangle = \left\langle -\left. \frac{\partial^2 l_1(\vec{\theta})}{\partial \theta_s \partial \theta_{s'}} \right|_{\vec{\theta}=\vec{\theta}^{(0)}} \right\rangle.$$

Таким образом на первом шаге получаем искомую величину:

$$\Delta\hat{\vec{\theta}}^{(1)} = [\tilde{C}^{(0)}]^{-1}\vec{d}^{(0)}, \quad \hat{\vec{\theta}}^{(1)} = \vec{\theta}^{(0)} + \Delta\hat{\vec{\theta}}^{(1)}.$$

На следующем шаге решаем систему линейных уравнений:

$$\Delta\hat{\vec{\theta}}^{(2)} = [\tilde{C}^{(1)}]^{-1}\vec{d}^{(1)}, \quad \hat{\vec{\theta}}^{(2)} = \hat{\vec{\theta}}^{(1)} + \Delta\hat{\vec{\theta}}^{(2)},$$

где

$$d_s^{(1)} = \left. \frac{\partial l_1(\vec{\theta})}{\partial \theta_s} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(1)}}, \quad \tilde{c}_{ss'}^{(1)} = \left\langle - \left. \frac{\partial^2 l_1(\vec{\theta})}{\partial \theta_s \partial \theta_{s'}} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(1)}} \right\rangle.$$

n -я итерация записывается в виде

$$\Delta\hat{\vec{\theta}}^{(n)} = [\tilde{C}^{(n-1)}]^{-1}\vec{d}^{(n-1)}, \quad \hat{\vec{\theta}}^{(n)} = \hat{\vec{\theta}}^{(n-1)} + \Delta\hat{\vec{\theta}}^{(n)},$$

где

$$d_s^{(n-1)} = \left. \frac{\partial l_1(\vec{\theta})}{\partial \theta_s} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n-1)}}, \quad \tilde{c}_{ss'}^{(n-1)} = \left\langle - \left. \frac{\partial^2 l_1(\vec{\theta})}{\partial \theta_s \partial \theta_{s'}} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n-1)}} \right\rangle.$$

Итерационная процедура заканчивается, когда выполняются пороговые условия

$$|\Delta\hat{\theta}_s^{(n)}/\hat{\theta}_s^{(n)}| \leq \beta \sim 10^{-2} \div 10^{-3}$$

для всех $s = 1, 2, \dots, S$.

В случае аддитивной модели

$$\vec{u} = \vec{f}(\vec{\theta}) + \vec{\varepsilon}$$

величины d_s и $\tilde{c}_{ss'}$ можно записать в явном виде

$$\begin{aligned} d_s^{(n)} &= (\vec{u} - \vec{f}(\vec{\theta}))^T R_\varepsilon^{-1} \left. \frac{\partial \vec{f}}{\partial \theta_s} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}}, \\ \tilde{c}_{ss'}^{(n)} &= \left\langle -(\vec{u} - \vec{f}(\vec{\theta}))^T R_\varepsilon^{-1} \left. \frac{\partial^2 \vec{f}}{\partial \theta_s \partial \theta_{s'}} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}} + \left. \frac{\partial \vec{f}^T}{\partial \theta_s} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}} R_\varepsilon^{-1} \left. \frac{\partial \vec{f}}{\partial \theta_{s'}} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}} \right\rangle = \\ &= - \underbrace{\langle \vec{\varepsilon}^T \rangle}_{=0} R_\varepsilon^{-1} \left. \frac{\partial^2 \vec{f}}{\partial \theta_s \partial \theta_{s'}} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}} + \left. \frac{\partial \vec{f}^T}{\partial \theta_s} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}} R_\varepsilon^{-1} \left. \frac{\partial \vec{f}}{\partial \theta_{s'}} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}} = \\ &= \left. \frac{\partial \vec{f}^T}{\partial \theta_s} R_\varepsilon^{-1} \frac{\partial \vec{f}}{\partial \theta_{s'}} \right|_{\vec{\theta}=\hat{\vec{\theta}}^{(n)}}. \end{aligned}$$

В качестве примеров рассмотрим оценку параметров намагниченных объектов.

Пример 1: Нить магнитных полюсов (вертикальная намагниченность)

$$f_k(\vec{\theta}) = \frac{2Mh}{h^2 + (x_k - \xi)^2}, \quad \vec{\theta} = \{M, h, \xi\},$$

где M – магнитный момент нити полюсов, h – глубина расположения нити полюсов, ξ – положение максимума функции $f_k(\vec{\theta})$.

Пример 2: Вертикальный магнитный диполь

$$f_k(\vec{\theta}) = \frac{M(2h^2 - x_k^2)}{(x_k^2 + h^2)^{5/2}}, \quad \vec{\theta} = \{M, h\},$$

где M – магнитный момент диполя, h – глубина расположения диполя.

Пример 3: Точечный магнитный полюс

$$f_k(\vec{\theta}) = \frac{Mh}{(x_k^2 + h^2)^{3/2}}, \quad \vec{\theta} = \{M, h\},$$

где M – магнитный момент, h – глубина расположения.

7.2 Метод наименьших квадратов: нелинейный случай

Рассмотрим модель наблюдений

$$\vec{u} = \vec{\varphi}(\vec{\theta}) + \vec{\varepsilon}.$$

После линеаризации функцию $\vec{\varphi}(\vec{\theta})$ в окрестности начальных значений параметров $\vec{\theta}^{(0)}$ можно представить в виде

$$\vec{\varphi}(\vec{\theta}) \approx \vec{\varphi}(\vec{\theta}^{(0)}) + \psi \Delta \vec{\theta},$$

где

$$\|\psi_{js}^{(0)}\| = \partial \varphi_j / \partial \theta_s |_{\vec{\theta} = \vec{\theta}^{(0)}} \quad \text{и} \quad \Delta \vec{\theta}^{(0)} = \vec{\theta} - \vec{\theta}^{(0)}.$$

Представим модель в виде

$$\tilde{u}_0 = \psi_0 \Delta \vec{\theta} + \vec{\varepsilon},$$

где $\tilde{u}_0 = \vec{u} - \vec{\varphi}(\vec{\theta}^{(0)})$.

Первую итерацию можно записать в явном виде

$$\hat{\Delta \vec{\theta}}^{(1)} = (\psi_0^T W_0 \psi_0)^{-1} \psi_0^T W_0 \tilde{u}_0,$$

где W_0 – весовая матрица:

$$W_0 = (\hat{\sigma}_{\varepsilon_0}^2)^{-1} I = (\hat{\varepsilon}_0^T \hat{\varepsilon}_0 / (n - S))^{-1} I_{S \times S},$$

при этом $\vec{\varepsilon}_0 = \vec{u} - \psi_0 \vec{\theta}^{(0)}$. Оценка вектора параметров, полученная на первом шаге:

$$\hat{\vec{\theta}}^{(1)} = \Delta \hat{\vec{\theta}}^{(1)} + \vec{\theta}^{(0)}.$$

Вторая итерация:

$$\hat{\vec{\theta}}^{(2)} = \Delta \hat{\vec{\theta}}^{(2)} + \hat{\vec{\theta}}^{(1)},$$

где

$$\begin{aligned}\Delta\hat{\theta}^{(2)} &= (\psi_1^T W_1 \psi_1)^{-1} \psi_1^T W_1 \tilde{u}_1, \\ \tilde{u}_1 &= \vec{u} - \varphi(\hat{\theta}^{(1)}), \quad \|\psi_{js}^{(1)}\| = \left. \frac{\partial \varphi_j}{\partial \theta_s} \right|_{\vec{\theta}=\hat{\theta}^{(1)}}, \\ W_1 &= (\hat{\sigma}_{\varepsilon_1}^2)^{-1} I = \left(\frac{\hat{\varepsilon}_1^T \hat{\varepsilon}_1}{n-S} \right)^{-1} I_{S \times S}, \\ \hat{\varepsilon}_1 &= \vec{u} - \psi_1 \hat{\theta}^{(1)}.\end{aligned}$$

Аналогично вычисляются поправки на n -м по результатам, полученным на $n-1$ -м шаге

$$\begin{aligned}\hat{\theta}^{(n)} &= \Delta\hat{\theta}^{(n)} + \hat{\theta}^{(n-1)}, \\ \Delta\hat{\theta}^{(n)} &= (\psi_{n-1}^T W_{n-1} \psi_{n-1})^{-1} \psi_{n-1}^T W_{n-1} \tilde{u}_{n-1}, \\ \tilde{u}_{n-1} &= \vec{u} - \varphi(\hat{\theta}^{(n-1)}), \quad \|\psi_{js}^{(n-1)}\| = \left. \frac{\partial \varphi_j}{\partial \theta_s} \right|_{\vec{\theta}=\hat{\theta}^{(n-1)}}, \\ W_{n-1} &= \left(\frac{\hat{\varepsilon}_{n-1}^T \hat{\varepsilon}_{n-1}}{n-S} \right)^{-1} I_{S \times S}, \\ \hat{\varepsilon}_{n-1} &= \vec{u} - \psi_{n-1} \hat{\theta}^{(n-1)}.\end{aligned}$$

Итерационный процесс завершается, если выполняются пороговые условия

$$|\Delta\hat{\theta}_s^{(n)} / \hat{\theta}_s^{(n)}| \leq \beta \sim 10^{-2} \div 10^{-3}$$

для всех $s = 1, 2, \dots, S$.

Следует отметить, что на каждом шаге итерационной процедуры мы используем метод наименьших квадратов применительно к линейной модели.

7.3 Метод наименьших квадратов для нестационарной модели

Рассмотрим модель с нестационарной случайной компонентой

$$\vec{u} = \psi \vec{\theta} + \vec{\varepsilon}.$$

Пусть дисперсия σ_i^2 , зависящая от времени

$$\sigma_i^2 = \Sigma(f_i, \vec{\rho}) = \rho_0 + \rho_1 f_i^2$$

пропорциональна энергии сигнала

$$f_i = \psi_i^T \vec{\theta}.$$

Задача состоит в оценке вектора параметров $\vec{\theta}$ и вектора параметров $\vec{\rho}(\rho_0, \rho_1)$ функции $\Sigma(f_i, \vec{\rho})$. Для решения задачи может быть использован следующий итерационный процесс.

1. В качестве начального вектора параметров $\vec{\theta}^{(0)}$ возьмем оценку, полученную (“обычным”) методом наименьших квадратов

$$\hat{\vec{\theta}}^{(0)} = (\psi^T \psi)^{-1} \psi^T \vec{u}.$$

2. Запишем оценку f_i :

$$\hat{f}_i = \vec{\psi}_i^T \hat{\vec{\theta}}^{(0)}.$$

3. Квадрат отклонения:

$$\hat{\varepsilon}_i^2 = (u_i - \hat{f}_i)^2.$$

Введем обозначение

$$\hat{B} = \begin{Bmatrix} 1 & \hat{f}_1^2 \\ 1 & \hat{f}_2^2 \\ \dots & \dots \\ 1 & \hat{f}_n^2 \end{Bmatrix}, \quad \hat{C} = \begin{Bmatrix} \hat{\varepsilon}_1^2 \\ \hat{\varepsilon}_2^2 \\ \dots \\ \hat{\varepsilon}_n^2 \end{Bmatrix}.$$

4. Оценка вектора параметров $\vec{\rho}$:

$$\hat{\vec{\rho}} = \arg \min (\hat{C} - \hat{B}\vec{\rho})^T (\hat{C} - \hat{B}\vec{\rho}), \quad \hat{\vec{\rho}} = (\hat{B}^T \hat{B})^{-1} \hat{B}^T \hat{C}.$$

5. Оценка дисперсии

$$\hat{\sigma}_i^2 = \hat{\rho}_0 + \hat{\rho}_1 \hat{f}_i^2.$$

6. Для оценивания вектора параметров $\vec{\theta}$ учтем нестационарность $\hat{\sigma}_i^{-2}$ введением весовой функции:

$$\tilde{\vec{\theta}} = (\psi^T W \psi)^{-1} \psi^T W \vec{u},$$

где

$$W = \text{diag} (\hat{\sigma}_1^{-2}, \hat{\sigma}_2^{-2}, \dots, \hat{\sigma}_n^{-2}).$$

7. Оценка квадрата расхождения:

$$\tilde{\varepsilon}_i^2 = (u_i - \tilde{f}_i)^2,$$

где

$$\tilde{f}_i = \vec{\psi}_i^T \tilde{\vec{\theta}}.$$

8. Итерационный процедура завершается, если выполнено пороговое условие

$$\tilde{\varepsilon}_i^2 \leq \beta_0.$$

Процесс продолжается, если $\tilde{\varepsilon}_i^2 \geq \beta_0$. В этом случае полагаем $\hat{\vec{\theta}}^{(0)} = \tilde{\vec{\theta}}$ и переходим к пункту 2.

7.4 Метод интервального оценивания

Пусть θ – неизвестный параметр. Точечная оценка параметра $\hat{\theta}$ является функцией результатов наблюдения

$$\hat{\theta} = \hat{\theta}(u_1, u_2, \dots, u_n).$$

Введем случайный интервал I_β , определяемый следующим образом

$$P(|\hat{\theta} - \theta| < \delta) = \beta, \quad I_\beta = [\hat{\theta} - \delta, \hat{\theta} + \delta],$$

где β – доверительная вероятность, I_β – доверительный интервал (см. рис. 7.1).

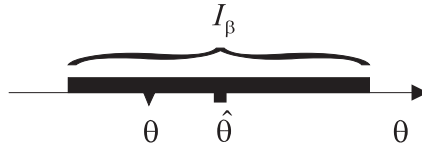


Рис. 7.1: Доверительный интервал I_β и оцениваемый параметр $\hat{\theta}$.

Пусть для модели $\vec{u} = \psi\vec{\theta} + \varepsilon$ получена МНК оценка $\hat{\theta}_s$, несмещенная $\langle \hat{\theta}_s \rangle = \theta_s$, с ковариационной матрицей

$$R_{\theta_s} = \sigma_\varepsilon^2 [(\psi^T W \psi)^{-1}]_{ss}, \quad s = 1, 2, \dots, S,$$

где W – весовая матрица.

После соответствующей нормировки получаем нормальное распределение:

$$\frac{\hat{\theta}_s - \theta_s}{\sigma_\varepsilon [(\psi^T W \psi)^{-1}]_{ss}^{1/2}} \in N(0, 1).$$

Для квадратичной формы отклонений (модельных и экспериментальных данных) имеем распределение χ^2 с $n - S$ степенями свободы:

$$\hat{\varepsilon} = \vec{u} - \psi\hat{\theta}, \quad \frac{\hat{\varepsilon}^T W \hat{\varepsilon}}{\sigma_\varepsilon^2} \in \chi_{n-S}^2.$$

Хорошо известно, что отношение случайной величины с нормальным распределением $N(0, 1)$ к величине, имеющей распределение χ^2 , имеет распределение Стьюдента (t -распределение):

$$t_{n-S} = \frac{\hat{\theta}_s - \theta_s}{\sigma_\varepsilon [(\psi^T W \psi)^{-1}]^{1/2} (1/\sigma_\varepsilon) [\hat{\varepsilon}^T W \hat{\varepsilon} / (n - S)]^{1/2}} \in \text{St}(t_{n-S})$$

Доверительная вероятность:

$$P[|t_{n-S}| \leq \gamma] = \beta.$$

Среднеквадратические отклонения $\hat{\sigma}_\varepsilon$ и $\hat{\sigma}_{\theta_s}$ определяются следующим образом

$$\hat{\sigma}_\varepsilon = \frac{\hat{\varepsilon}^T W \hat{\varepsilon}}{n - S}, \quad \hat{\sigma}_{\theta_s} = \hat{\sigma}_\varepsilon [(\psi^T W \psi)^{-1}]_{ss}.$$

Доверительный интервал для параметра θ_s записывается в виде

$$I_\beta = \left[\hat{\theta}_s \pm \gamma [(\psi^T W \psi)^{-1}]_{ss}^{1/2} \left[\frac{\hat{\varepsilon}^T W \hat{\varepsilon}}{n - S} \right]^{1/2} \right] = [\hat{\theta}_s \pm \gamma \hat{\sigma}_{\theta_s}].$$

В случае оценки для функции f :

$$\vec{f} = \psi \vec{\theta}$$

При нахождении доверительного интервала будем учитывать

$$\frac{\hat{f}_i - f_i}{\sigma_\varepsilon [\psi (\psi^T W \psi)^{-1} \psi^T]_{ii}^{1/2}} \in N(0, 1)$$

и

$$t_{n-S} = \frac{\hat{f}_i - f_i}{[\psi (\psi^T W \psi)^{-1} \psi^T]_{ii}^{1/2} [\hat{\varepsilon}^T W \hat{\varepsilon} / (n - S)]^{1/2}} \in \text{St}(t_{n-S}).$$

При этом доверительная вероятность запишется следующим образом

$$P(|t_{n-S}| \leq \gamma) = \beta,$$

а доверительная область дается выражением

$$I_\beta = \hat{f}_i \pm \gamma \hat{\sigma}_\varepsilon [\psi (\psi^T W \psi)^{-1} \psi^T]_{ii}^{1/2}$$

(см. рис. 7.2).

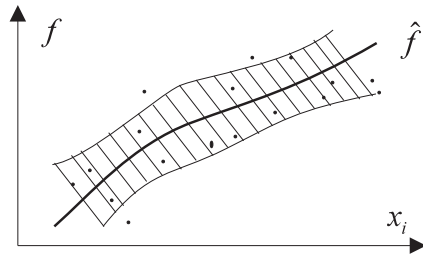


Рис. 7.2: Графическое представление доверительной области I_β .

7.5 Генетический алгоритм

Генетические алгоритмы (ГА), впервые предложенные Джоном Голландом (1975), основаны на аналогии с процессом биологической эволюции. Генетические алгоритмы демонстрируют высокую эффективность при решении нелинейных задач. Основные элементы ГА следующие: *кодирование*, *отбор*, *скрещивание*, *мутация*.

7.5.1 Кодирование

Общей чертой всех ГА является двоичная кодировка параметров модели.

Рассмотрим, например, кодирование скорости волн сжатия. Пусть нижний предел скорости полагается равным 1500 м/с, а верхний предел – 1810 м/с. Предположим, что требуемое разрешение – 10 м/с. Таким образом для подобного кодирования достаточно пяти бит (см. рис. 7.3). На этом этапе алгоритм

0	0	0	0	0	$v_{min} = 1500$ м/с
0	0	0	0	1	$v = 1510$ м/с
0	0	0	1	0	$v = 1520$ м/с
0	0	0	1	1	$v = 1530$ м/с
.....				
1	1	1	1	1	$v_{max} = 1810$ м/с

Рис. 7.3: Двоичное кодирование величин скорости.

определяет пригодность модели. Это означает, что двоичная информация кодируется с учетом физической модели, т.е. решается прямая задача. Полученные синтетические данные сравниваются с действительными наблюдаемыми данными, при этом используется специальный критерий согласия. В зависимости от решаемой задачи можно использовать различные критерии согласия. Например, рассмотрим нормированную корреляционную функцию

$$F(\vec{\theta}) = \frac{u_0 \otimes u_s(\vec{\theta})}{(u_0 \otimes u_0)^{1/2} (u_s \otimes u_s(\vec{\theta}))^{1/2}}$$

и среднеквадратичную функцию ошибки

$$F(\vec{\theta}) = \sum_{i=1}^n (u_{i0} - u_{is}(\vec{\theta}))^2,$$

где \otimes – обозначает операцию вычисления корреляционной функции, u_0 и $u_s(\vec{\theta})$ соответствуют наблюдаемым данным и модельным данным для параметра $\vec{\theta}$.

7.5.2 Отбор

Отбор пар моделей для воспроизводства основан на величинах согласия между ними. Модели с высокими величинами согласия с большей вероятностью подвержены отбору, чем модели с низкими величинами согласия. Рассмотрим отбор, пропорциональный согласию моделей.

Наиболее часто используются методы отбора, основанные на отношении функций согласия индивидуальных моделей к сумме согласий для всех моделей из

рассматриваемого набора (популяции). Таким образом оценивается вероятность их отбора:

$$p_s(\theta_i) = \frac{F(\theta_i)}{\sum_{j=1}^n F(\theta_j)},$$

где n – число рассматриваемых моделей. Отбор, базирующийся на подобных вероятностях, продолжается до тех пор, пока множество моделей не будет разбито на пары.

В типичных ГА, при первоначальном наборе, содержащем 100 моделей, 50 пар отбирают на основе их величин согласия. Рассмотрим модель, которая удовлетворяет равенству

$$p_s(\theta_i) \geq \delta,$$

где δ – заданная пороговая величина. Пусть имеется L моделей, которые формируют $L/2$ пар. Каждая пара моделей теперь порождает две вторичные пары при использовании генетических операций скрещивания и мутации. Как результат образуется новая популяция.

7.5.3 Скрещивание

Скрещивание представляет собой механизм, позволяющий разделить генетическую информацию между парными моделями. Если использовать терминологию геофизических обратных задач, то скрещивание позволяет производить обмен информацией между парными моделями и как следствие порождать новые модели. Предполагается, что скрещивание может быть одноточечным и многоточечным. В случае одноточечного скрещивания простая (однобитовая) позиция двоичного набора выбирается случайно в соответствии с равномерным распределением. Все двоичные состояния, расположенные справа от заданного состояния, осуществляют обмен между двумя моделями, генерируя две новые модели (см. рис. 7.4). В случае многоточечного скрещивания, эта операция выполняется независимо для каждого параметра модели (см. рис. 7.5).

Пример. Пара параметров скоростной модели v_1 и v_2 являются экстремальными членами для схемы кодирования от 1500 м/с до 1810 м/с (см. 7.6 и табл. 7.1).

Crossover position	v_1 m/s	v_2 m/s
0	1500	1810
1	1510	1800
2	1530	1780
3	1570	1740
4	1650	1660
5	1810	1500

Таблица 7.1: Пара экстремальных положений для схемы кодирования.

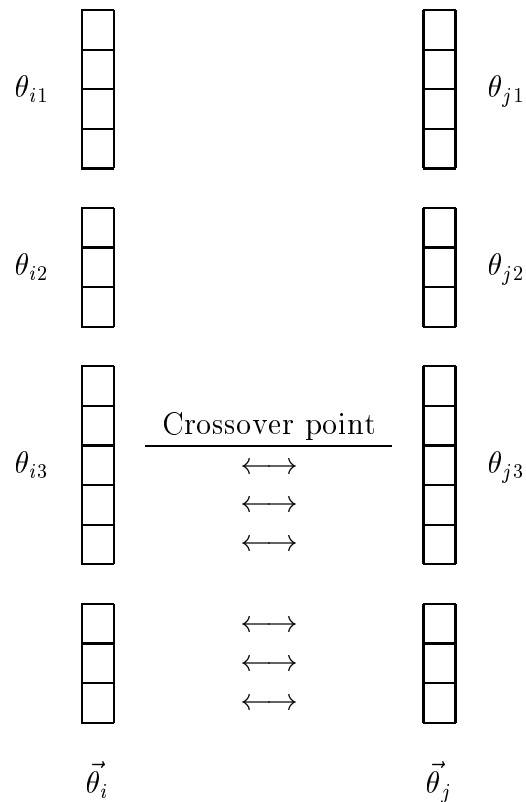


Рис. 7.4: Пример одноточечного скрещивания.

7.5.4 Мутация

Мутация – последняя генетическая операция. Мутация заключается в случайном изменении двоичного состояния. Такая операция может быть выполнена на этапе скрещивания. Возможность выполнения мутации в том или ином случае конкретном случае обычно задается в ГА в вероятностной форме.

Пример. Процесс мутации для начальной величины скорости 1760 м/с (см. рис. 7.7).

Если величина ε , полученная при помощи случайного генератора, меньше чем пороговое значение P_m , то процедура мутации выполняется, если $\varepsilon > P_m$, то процедура мутации не выполняется.

7.5.5 Выбор

Выбор осуществляется для каждой из $L/2$ пар модели, которые имеют наибольшие функции согласия (см. рис. 7.8). Таким образом, получены $L/2$ моделей, которые случайным образом формируют $L/4$ пары. После этого повторяются процедуры: скрещивание, мутация, выбор. Эта процедура повторяется до тех пор, пока не будет получена оптимальная модель. Основным преимуществом ГА

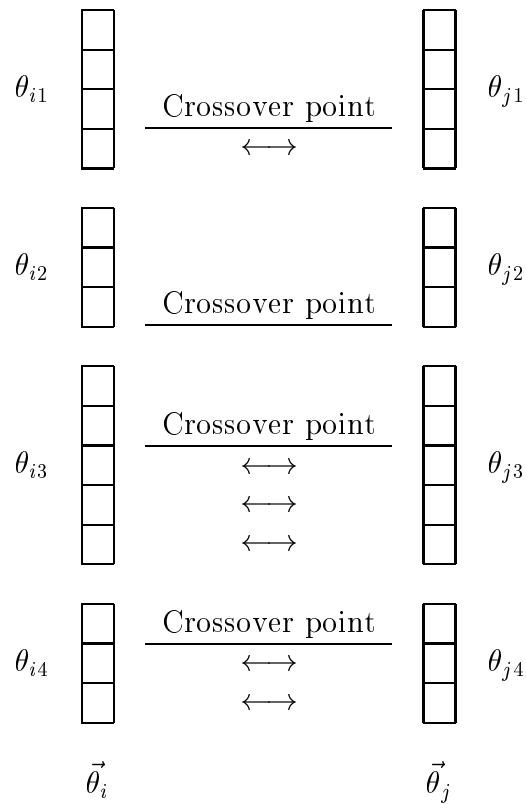


Рис. 7.5: Пример многоточечного скрещивания.

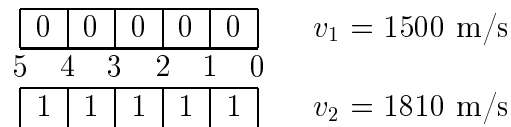


Рис. 7.6: Пара экстремальных положений для схемы кодирования.

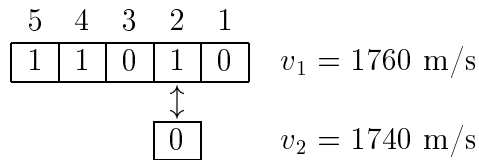


Рис. 7.7: Иллюстрация процедуры мутации.

является возможность нахождения глобального максимума с существенно меньшими затратами времени, чем, например, в случае метода Метода-Карло. Отметим, что метод локальной линеаризации позволяет находить только локальный экстремум.

Пример 1: Тест ГА на синтетической сейсмограмме.

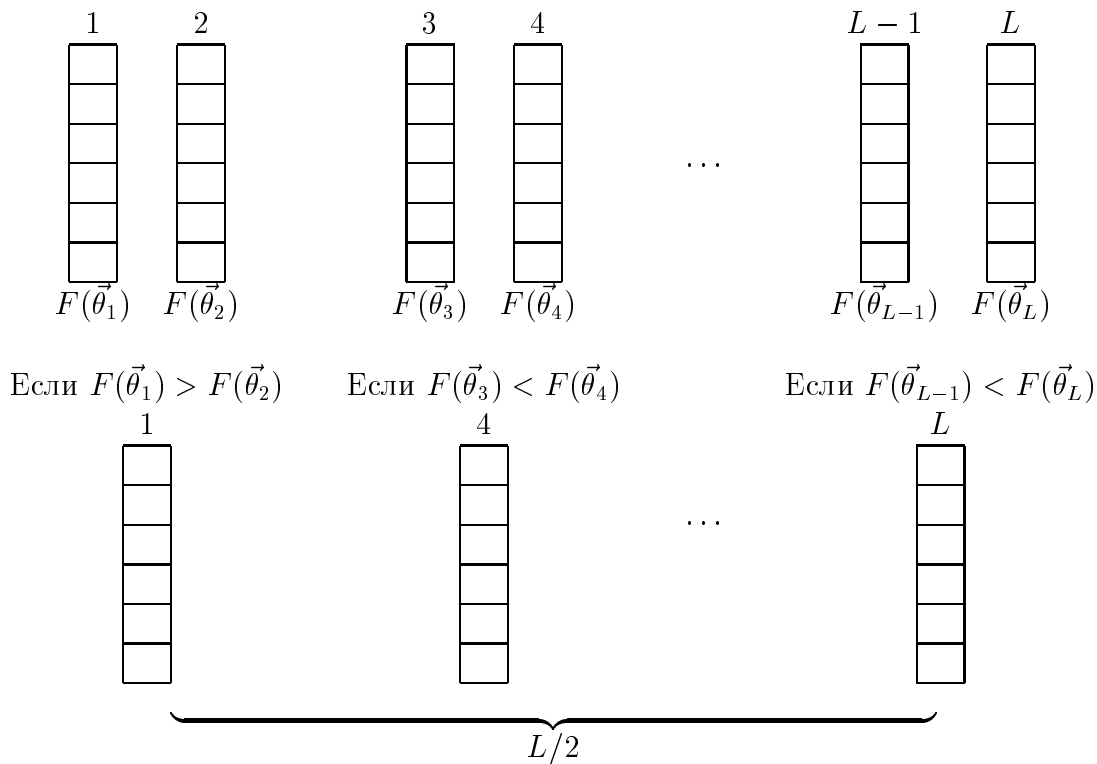


Рис. 7.8: А Изменение модели в процессе реализации ГА.

В случае сейсмического сигнала функция согласия может иметь вид

$$F(A_1, \dots, A_M, \tau_1, \dots, \tau_M) = \sum_{i=1}^n \left(u_i - \sum_{\mu=1}^M A_\mu \varphi(t_i - \tau_\mu) \right)^2,$$

для формы сейсмического сигнала

$$\varphi = \exp\{-\beta|t|\} \cos \omega t$$

где A_μ – амплитуда, τ_μ – время прихода сейсмической волны с номером μ . Возможные границы изменения неизвестных параметров (A_μ , τ_μ) и дискретного интервала задаются из априорных соображений. Параметр затухания β и частота ω фиксированы. Оценки параметров находятся из условия минимизации функции согласия $F(A_1, \dots, A_M, \tau_1, \dots, \tau_M)$.

Применяя ГА, мы выбирали тип кодирования битовой последовательности, размер рабочей популяции L и величины вероятностей скрещивания и мутации.

Пример 2: Пусть размер популяции, вероятности скрещивания и мутации равны соответственно:

$$L = 50, \quad P_c = 0.6, \quad P_m = 0.001.$$

Алгоритм тестируется при различных размерах популяции, отношениях мутации и скрещивания.

В случае двух интерферирующих сейсмических волн пороговое условия правильного разделения волн – $1/4$ периода волны, при этом отношение сигнал/помеха должно быть больше или равно двум. Рассмотрим три интерферирующие сейсмические волны при условии $|\tau_1 - \tau_2| = |\tau_2 - \tau_3| \geq 0.016$ с. Три волны разделены значимыми значениями оцениваемых параметров, которые соответствуют величине временной задержки между волнами в пределах $1/4$ и $1/3$ периода волны. Отношение сигнал/шум должно быть также больше или равно двум.

Глава 8

Статистические критерии выбора модели

Обратные геофизические задачи не обладают свойством единственности. В процессе интерпретации рассматривают некоторое множество конкурирующих моделей. Решение обратной задачи находят с учетом геологической и геофизической априорной информации относительно каждой модели. При этом используют статистические критерии выбора модели.

8.1 Параметрические критерии выбора модели

Пусть дан вектор параметров $\vec{\theta}_0$. Гипотеза H_0 состоит в том, что вектор параметров $\vec{\theta}$ равен $\vec{\theta}_0$, а гипотеза H_1 состоит в том, что вектор параметров $\vec{\theta}$ не равен $\vec{\theta}_0$:

$$\begin{aligned} H_0 &: \vec{\theta} = \vec{\theta}_0, \\ H_1 &: \vec{\theta} \neq \vec{\theta}_0. \end{aligned}$$

Используем в качестве критерия отношение правдоподобия

$$\lambda = -2 \ln \frac{L(\vec{u}, \vec{\theta}_0)}{L(\vec{u}, \vec{\theta})} \in \chi_S^2.$$

Отношение правдоподобия λ принадлежит распределению χ_S^2 с S степенями свободы. Критерий может быть выбран в виде

$$\begin{cases} \lambda < \chi_{\alpha, S}^2 & \text{то гипотеза } H_0 \text{ не противоречит данным } \vec{u}, \\ \lambda > \chi_{\alpha, S}^2 & \text{то гипотеза } H_0 \text{ отвергается.} \end{cases}$$

Пример: Рассмотрим модель

$$\vec{u} = \begin{cases} \text{или } \vec{f}(\vec{\theta}_0) + \vec{\varepsilon}_0 & : H_0 \quad \vec{\varepsilon}_0 \in N(0, \sigma_0^2 I), \\ \text{или } \vec{f}(\vec{\theta}_1) + \vec{\varepsilon}_1 & : H_1 \quad \vec{\varepsilon}_1 \in N(0, \sigma_1^2 I). \end{cases}$$

Отношение правдоподобия имеет вид

$$\lambda = -2 \ln \left[\frac{L(\vec{u}, \vec{\theta}_0, \sigma_0^2)}{L(\vec{u}, \vec{\theta}_1, \sigma_1^2)} \right].$$

Оценки дисперсий:

$$\hat{\sigma}_0^2 = (1/n)(\vec{u} - \vec{f}(\vec{\theta}_0))^2, \quad \hat{\sigma}_1^2 = (1/n)(\vec{u} - \vec{f}(\vec{\theta}_1))^2.$$

Подставляя эти оценки в отношение λ -отношение, получаем

$$\lambda = n \ln(\hat{\sigma}_0^2 / \hat{\sigma}_1^2).$$

8.2 Критерий отношения апостериорных вероятностей

Рассмотрим модель выбора между двумя сигналами $\vec{f}(\vec{\theta}_0)$ и $\vec{f}(\vec{\theta}_1)$

$$\vec{u} = \begin{cases} \text{или } \vec{f}(\vec{\theta}_0) + \vec{\varepsilon} & : H_0, \\ \text{или } \vec{f}(\vec{\theta}_1) + \vec{\varepsilon} & : H_1. \end{cases}$$

Критерий апостериорных вероятностей можно представить в виде

$$\lambda = \ln \frac{P(1|\vec{u})}{P(0|\vec{u})} = \ln \frac{P(1)p(\vec{u}/1)}{P(0)p(\vec{u}/0)},$$

где $P(0)$ и $P(1)$ – априорная информация соответственно относительно гипотез H_0 и H_1 . Применение критерия сводится к проверке

$$\begin{cases} \text{Если } \lambda < 0 & \text{то гипотеза } H_0 \text{ не противоречит данным } \vec{u}, \\ \text{Если } \lambda \geq 0 & \text{то гипотеза } H_0 \text{ отвергается.} \end{cases}$$

Предполагаем, что $\vec{\varepsilon} \in N(0, R)$, то

$$\lambda = \frac{1}{2}(\vec{u} - \vec{f}(\vec{\theta}_0))^T R^{-1}(\vec{u} - \vec{f}(\vec{\theta}_0)) - \frac{1}{2}(\vec{u} - \vec{f}(\vec{\theta}_1))^T R^{-1}(\vec{u} - \vec{f}(\vec{\theta}_1)). \quad (8.1)$$

Найдем условные математические ожидания и дисперсии:

если гипотеза H_0 верна и $\vec{u} = \vec{f}(\vec{\theta}_0) + \vec{\varepsilon}$, то будем находить $\langle \lambda_0 \rangle$ и $\sigma_{\lambda_0}^2$;

если гипотеза H_1 верна и $\vec{u} = \vec{f}(\vec{\theta}_1) + \vec{\varepsilon}$, то будем находить $\langle \lambda_1 \rangle$ и $\sigma_{\lambda_1}^2$.

Для нахождения $\langle \lambda_0 \rangle$ подставим в (8.1) вместо \vec{u} модель $\vec{f}(\vec{\theta}_0) + \vec{\varepsilon}$

$$\begin{aligned} \langle \lambda_0 \rangle &= \frac{1}{2} \langle \vec{\varepsilon}^T R^{-1} \vec{\varepsilon} \rangle - \frac{1}{2} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1)) - \\ &- \frac{1}{2} \langle \vec{\varepsilon}^T R^{-1} \vec{\varepsilon} \rangle - \langle (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} \vec{\varepsilon} \rangle = \\ &= -\frac{1}{2} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1)). \end{aligned}$$

Легко проверить, подставляя вместо \vec{u} модель $\vec{f}(\vec{\theta}_1) + \vec{\varepsilon}$, что

$$\langle \lambda_1 \rangle = -\langle \lambda_0 \rangle.$$

Найдем $\sigma_{\lambda_0}^2$, используя вместо \vec{u} модель $\vec{f}(\vec{\theta}_0) + \vec{\varepsilon}$

$$\begin{aligned} \sigma_{\lambda_0}^2 &= \langle (\lambda_0 - \langle \lambda_0 \rangle)^2 \rangle = \langle \left(-\frac{1}{2}(\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))\right)^T R^{-1} \times \\ &\times (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1)) - (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} \vec{\varepsilon} + \frac{1}{2}(\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T \times \\ &\times R^{-1}(\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1)) \rangle = \\ &= (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} \langle \vec{\varepsilon} \vec{\varepsilon}^T \rangle R^{-1} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1)) = \\ &= (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1)). \end{aligned}$$

Легко доказать, что

$$\sigma_{\lambda_1}^2 = \sigma_{\lambda_0}^2.$$

Таким образом мы можем записать условные вероятности для λ

$$p(\lambda/0) = \frac{1}{\sqrt{2\pi}\sigma_{\lambda_0}} \exp\left\{-\frac{(\lambda - \lambda_0)^2}{2\sigma_{\lambda_0}^2}\right\},$$

при условии, что гипотеза H_0 верна и

$$p(\lambda/1) = \frac{1}{\sqrt{2\pi}\sigma_{\lambda_1}} \exp\left\{-\frac{(\lambda - \lambda_1)^2}{2\sigma_{\lambda_1}^2}\right\},$$

при условии, что гипотеза H_1 верна.

Рассмотрим четыре случая, возникающих при принятии решения (см. рис. 8.1):

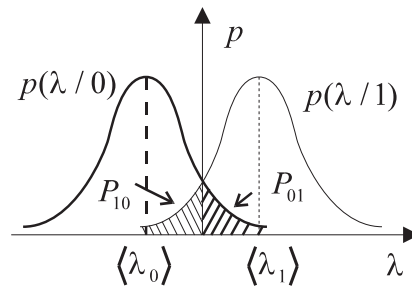


Рис. 8.1: Условные плотности для гипотез H_0 и H_1 и ошибки принятия решений.

1. Гипотеза H_0 верна и она не отвергается согласно критерию.

2. Гипотеза H_0 верна, но она отвергается (ошибка первого рода). Вероятность P_{01} есть интеграл от $p(\lambda/0)$ в интервале от 0 до $+\infty$

$$\begin{aligned} P_{01} &= \int_0^{\infty} p(\lambda/0) d\lambda = \frac{1}{\sqrt{2\pi}\sigma_{\lambda_0}} \int_0^{\infty} \exp\left\{-\frac{(\lambda - \langle\lambda_0\rangle)^2}{2\sigma_{\lambda_0}^2}\right\} d\lambda = \\ &= \left[1 - \Phi\left(-\frac{\langle\lambda_0\rangle}{\sigma_{\lambda_0}}\right)\right], \end{aligned}$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{t^2}{2}\right\} dt$$

– функция Лапласа ($\Phi(-x) = 1 - \Phi(x)$).

3. Гипотеза H_0 неверна и она отвергается согласно критерию.
4. Гипотеза H_0 неверна, но принимается согласно критерию (ошибка второго рода). Вероятность ошибки P_{10} есть интеграл от $p(\lambda/1)$ в интервале от 0 до $-\infty$.

$$\begin{aligned} P_{10} &= \int_{-\infty}^0 p(\lambda/1) d\lambda = \frac{1}{\sqrt{2\pi}\sigma_{\lambda_1}} \int_{-\infty}^0 \exp\left\{-\frac{(\lambda - \langle\lambda_1\rangle)^2}{2\sigma_{\lambda_1}^2}\right\} d\lambda = \\ &= \Phi\left(-\frac{\langle\lambda_1\rangle}{\sigma_{\lambda_1}}\right), \end{aligned}$$

Принимая во внимание $\sigma_{\lambda_0} = \sigma_{\lambda_1}$ и $\langle\lambda_0\rangle = -\langle\lambda_1\rangle$ – получим выражение для полной вероятности ошибки

$$P_e = P(0)P_{01} + P(1)P_{10}.$$

В случае $P(0) = P(1) = 1/2$

$$\begin{aligned} P_e &= \frac{1}{2} \left[1 - \Phi\left(-\frac{\langle\lambda_0\rangle}{\sigma_{\lambda_0}}\right)\right] + \frac{1}{2} \Phi\left(\frac{\langle\lambda_0\rangle}{\sigma_{\lambda_0}}\right) = \Phi\left(\frac{\langle\lambda_0\rangle}{\sigma_{\lambda_0}}\right), \\ P_e &= \Phi(-\alpha/2) = 1 - \Phi(\alpha/2), \\ \alpha &= [(\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))]^{1/2}. \end{aligned}$$

Графическое представление $\Phi(-\alpha/2)$ – см. рис. 8.2.

Запишем пороговое условие для полной вероятности 5 %:

$$[(\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))^T R^{-1} (\vec{f}(\vec{\theta}_0) - \vec{f}(\vec{\theta}_1))]^{1/2} \geq \alpha_0 = 3.3.$$

Пример: Рассмотрим задачу обнаружения сейсмических волн. Пусть модель некоторой части сейсмограммы допускает представление в виде

$$u_{ki} = \begin{cases} \text{или} & A\varphi(t_i - \tau - k\Delta x\gamma) + \varepsilon_{ki} & : H_0, \\ \text{или} & \varepsilon_{ki} & : H_1, \end{cases}$$

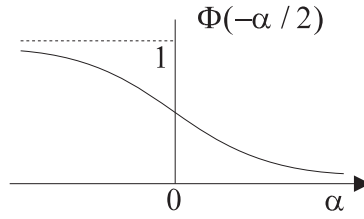
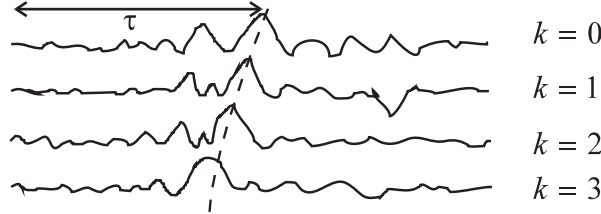
Рис. 8.2: Графическое представление $\Phi(-\alpha/2)$.

Рис. 8.3: Иллюстрация обнаружения сейсмического сигнала.

где A – амплитуда, τ – время прихода на трассе $k = 0$, γ – кажущаяся скорость сейсмической волны, ε_{ki} – случайная ошибка, $\varphi(t)$ – форма сейсмического сигнала, $\varepsilon_{ki} \in N(0, R)$ (см. рис. 8.3).

Если полная вероятность ошибки равна 5 %, то пороговое условие имеет вид

$$\alpha^2 = \frac{A^2}{\sigma^2} \sum_k \sum_i \sum_{i'} r_{ii'}^{-1} \varphi(t_i - \tau - k\Delta x\gamma) \varphi(t_{i'} - \tau - k\Delta x\gamma) \geq 3.3^2.$$

В случае некоррелированного шума $R = \sigma^2 I$:

$$\alpha^2 = \frac{A^2}{\sigma^2} \sum_{i=1}^n \sum_{k=1}^K \varphi^2(t_i - \tau - k\Delta x\gamma) \geq 3.3^2.$$

Если такое условие выполняется, то говорят, что сейсмическая волна с параметрами (A, τ, γ) существует с ошибкой (ненадежностью) 5 %.

8.3 Задача разделения сигналов

Рассмотрим модель для разделения сигналов

$$u = \begin{cases} \text{или } \vec{f}(\vec{\theta}_1) + \vec{f}(\vec{\theta}_2) + \vec{\varepsilon} & : H_0, \\ \text{или } \vec{f}(\vec{\theta}_0) + \vec{\varepsilon} & : H_1. \end{cases}$$

Гипотеза H_0 : имеются два сигнала, альтернативная гипотеза H_1 имеется один сигнал. Пусть $\varepsilon \in N(0, R)$, что приводит к пороговому условию. Если

$$\alpha^2 = (\vec{f}(\vec{\theta}_1) + \vec{f}(\vec{\theta}_2) - \vec{f}(\vec{\theta}_0))^T R^{-1} (\vec{f}(\vec{\theta}_1) + \vec{f}(\vec{\theta}_2) - \vec{f}(\vec{\theta}_0)) \geq \alpha_{0Pe}^2,$$

то можно утверждать, что с полной вероятностью ошибки P_e , мы имеем два сигнала.

Пример 1: Рассмотрим разделение сейсмических сигналов. Запишем модель в виде

$$u_{ki} = \begin{cases} \text{или} & A_1\varphi(t_i - \tau_1 - k\Delta x\gamma_1) + A_2\varphi(t_i - \tau_2 - k\Delta x\gamma_2) + \varepsilon_{ki} & : H_0, \\ \text{или} & A_0\varphi(t_i - \tau_0 - k\Delta x\gamma_0) + \varepsilon_{ki} & : H_1. \end{cases}$$

Функция α^2 имеет следующий вид

$$\alpha^2 = \frac{1}{\sigma^2} \sum_k \sum_i \sum_{i'} r_{ii'}^{-1} (A_1\varphi(t_i - \tau_1 - k\Delta x\gamma_1) + A_2\varphi(t_i - \tau_2 - k\Delta x\gamma_2) - A_0\varphi(t_i - \tau_0 - k\Delta x\gamma_0)) (A_1\varphi(t_{i'} - \tau_1 - k\Delta x\gamma_1) + A_2\varphi(t_{i'} - \tau_2 - k\Delta x\gamma_2) - A_0\varphi(t_{i'} - \tau_0 - k\Delta x\gamma_0)).$$

Введем упрощающие предположения

$\tau_1 = \tau_0 - \Delta\tau$, $\gamma_1 = \gamma_0 - \Delta\gamma$, $\tau_2 = \tau_0 + \Delta\tau$, $\gamma_2 = \gamma_0 + \Delta\gamma$, $A_0 = A_1/2 = A_2/2$ и $R = \sigma^2 I$ (некоррелированный шум). С учетом сделанных предположений получим:

$$\alpha^2 = \frac{A_0^2}{\sigma^2} \sum_k \sum_i \left[\frac{1}{2} \varphi(t_i - (\tau_0 - \Delta\tau) - k\Delta x(\gamma_0 - \Delta\gamma)) + \frac{1}{2} \varphi(t_i - (\tau_0 + \Delta\tau) - k\Delta x(\gamma_0 + \Delta\gamma)) - \varphi(t_i - \tau_0 - k\Delta x\gamma_0) \right]^2.$$

В частном случае $\Delta\gamma = 0$ функция α^2 от $\Delta\tau$ имеет вид – см. рис. 8.4).

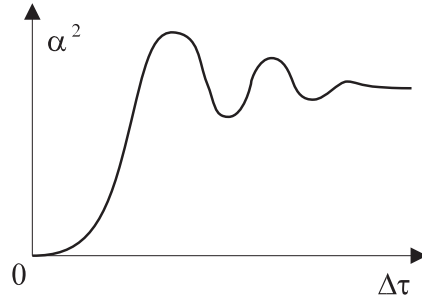


Рис. 8.4: Графическое представление функции α^2 .

Пример 2: Разделение магнитных полей от двух вертикальных тонких слоев с бесконечно погруженными нижними кромками

Разделяемые сигналы запишем в виде

$$\begin{aligned} f_{1k} &= \frac{I}{h} \left[\frac{1}{1 + (k\Delta x/h - \xi/h)^2} \right], & f_{2k} &= \frac{I}{h} \left[\frac{1}{1 + (k\Delta x/h + \xi/h)^2} \right], \\ f_{0k} &= \frac{2I}{h} \left[\frac{1}{1 + (k\Delta x/h)^2} \right], & & (8.2) \end{aligned}$$

где I – намагниченность, ξ – горизонтальная координата слоя, h – глубина нижней кромки. Сумма функций: f_{0k} , f_{1k} и f_{2k} ($f = f_{0k} + f_{1k} + f_{2k}$) представлены на рис. 8.5. Запишем функцию α^2 и пороговое условие

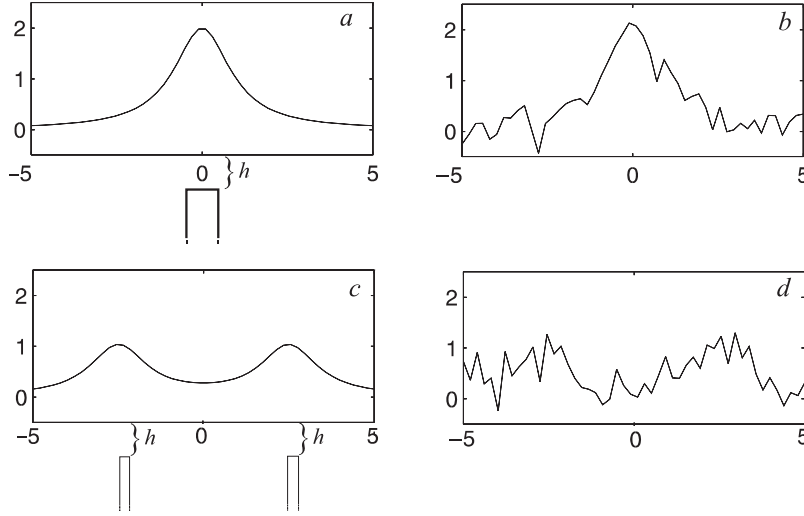


Рис. 8.5: Магнитное поле от тонких слоев определяемых выражениями из (8.2) ($I = 1$, $\xi = 2.5$, $h = 1.0$). а, б – без шума; в, г – с некоррелированным нормальным шумом $N(0, 0.1)$.

$$\left(\frac{2I}{\sigma h}\right)^2 \sum_{k=-K}^K \left[\left(2\left(1 + \left(\frac{k\Delta x}{h} - \frac{\xi}{h}\right)^2\right)\right)^{-1} + \left(2\left(1 + \left(\frac{k\Delta x}{h} + \frac{\xi}{h}\right)^2\right)\right)^{-1} - \left(1 + \left(\frac{k\Delta x}{h}\right)^2\right)^{-1} \right]^2 \geq \alpha_{0Pe}^2.$$

При достаточно большом h , если неравенство

$$\left(\frac{k\Delta x}{h} + \frac{\xi}{h}\right)^2 < 0.5$$

выполнено, можно записать приближенное выражение для порогового условия

$$\left(\frac{\xi}{h}\right)^2 \geq \frac{\sigma h}{2I} \frac{\alpha_0}{\sqrt{2k+1}}.$$

Пороговое условие и его приближенный аналог определяют величину ξ при которой еще возможно отличить интерференционное поле для двух слоев от поля одиночного слоя.

Пример 3: Проиллюстрируем применение метода максимального правдоподобия и метода отношения апостериорных вероятностей к оценке параметра сигнала интерферирующих полей.

Модель измерений может быть записана в виде

$$\vec{u}_k = \sum_{\mu=1}^M \vec{f}_{\mu k}(\vec{\theta}_\mu) + \vec{\varepsilon}_k, \quad \vec{\varepsilon}_k \in N(0, R).$$

В сейсмическом случае вектор $\vec{f}_{\mu k}$ имеет вид

$$\vec{f}_{\mu k} = \|A_\mu \varphi(t_i - \tau_\mu - k\Delta x \gamma_\mu)\|_{i=1}^n, \quad \vec{\theta}_\mu = \{A_\mu, \tau_\mu, \gamma_\mu\}, \quad k = 0, 1, \dots, K,$$

где A_μ – амплитуда, τ_μ – время прихода на трассе с номером $k = 0$, γ_μ – обратная кажущаяся скорость распространения сейсмического сигнала с номером μ . Пример интерференции сейсмических сигналов представлен на рис. 8.6

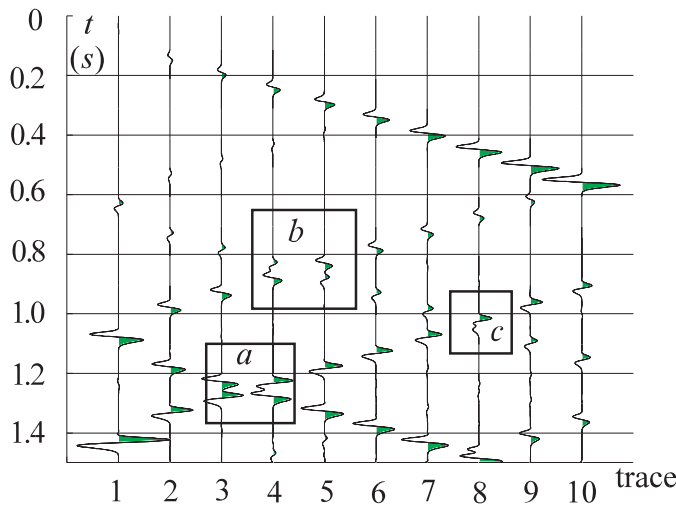


Рис. 8.6: Иллюстрация интерференции сейсмических волн. Модель среды та же, что на рис. 5.1. Точки наблюдения: 0.003, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5; a, b, c – области с интерференцией сейсмических волн. (Сейсмограмма изображена с автоматическим усилением.)

Рассмотрим случай интерференции магнитных полей. Интерференционное магнитное поле от неизвестного числа вертикальных тонких слоев с бесконечно погруженной нижней кромкой имеет вид

$$\vec{u}_k = \sum_{\mu=1}^M \vec{f}_k(\vec{\theta}_\mu) + \vec{\varepsilon}_k, \quad \vec{\varepsilon}_k \in N(0, R),$$

$$f_k(\theta_\mu) = \frac{I_\mu h_\mu}{h_\mu^2 + (k\Delta x - \xi_\mu)^2}, \quad \vec{\theta}_\mu = \{I_\mu, h_\mu, \xi_\mu\},$$

где I_μ – намагниченность слоя, ξ_μ – горизонтальная координата слоя, h_μ – глубина расположения нижней кромки с номером μ . Пример магнитного поля от шести вертикальных слоев представлен на рис. 8.7.

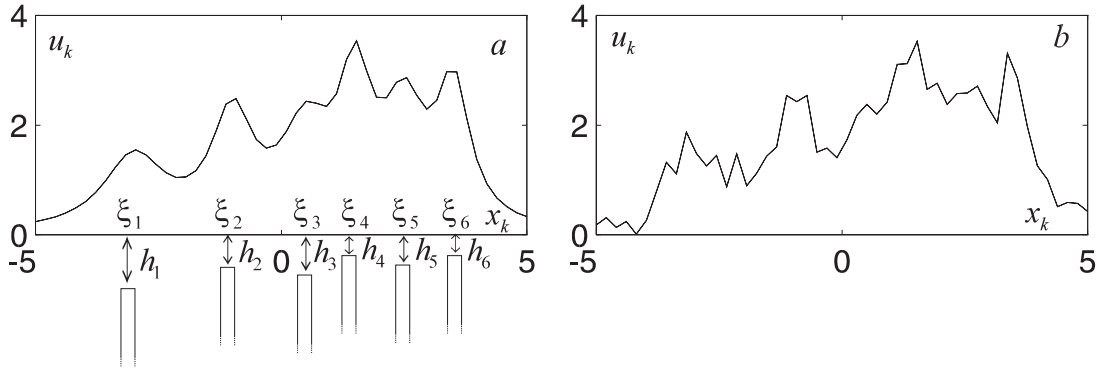


Рис. 8.7: Магнитное поле вызванное шестью тонкими слоями ($I_\mu = 1$). Пары параметров (h_μ, ξ_μ) ($\mu = 1, \dots, 6$) – следующие: $(0.75, -3.0)$, $(0.5, -1.0)$, $(0.6, 0.5)$, $(0.5, 1.5)$, $(0.5, 2.5)$, $(0.4, 3.5)$. а – без шума; б – с нормальным некоррелированным шумом $N(0, 0.2)$.

Задача состоит в получении оценки для вектора параметров $\vec{\theta}_\mu$, ($\mu = 1, 2, \dots, M$) и числа источников поля M

$$\vec{\theta}_\mu - ? \quad M - ?$$

С учетом сделанных предположений решающая функция, полученная на основе метода максимального правдоподобия записывается следующим образом

$$l(\vec{u}_k, \vec{\theta}) = -\frac{1}{2} \sum_k (\vec{u}_k - \sum_{\mu=1}^M \vec{f}_{\mu k}(\vec{\theta}_\mu))^T R^{-1} (\vec{u}_k - \sum_{\mu=1}^M \vec{f}_{\mu k}(\vec{\theta}_\mu)).$$

Оценка параметров в общем виде представляется формулой

$$\hat{\vec{\theta}} = \arg \max_{\vec{\theta}} l(\vec{u}_k, \vec{\theta}), \quad \vec{\theta} = \|\vec{\theta}_1, \vec{\theta}_2, \dots, \vec{\theta}_M\|.$$

Если в функции $l(\vec{u}_k, \vec{\theta})$ оставить только члены, которые зависят от вектора параметров $\vec{\theta}$ то получим

$$g(\vec{u}_k, \vec{\theta}) = \sum_{\mu=1}^M \sum_k [\vec{u}_k^T R^{-1} \vec{f}_{\mu k}(\vec{\theta}_\mu) - \frac{1}{2} \sum_{\mu' \neq \mu} \vec{f}_{\mu' k}^T(\vec{\theta}_{\mu'}) R^{-1} \vec{f}_{\mu k}(\vec{\theta}_\mu) - \frac{1}{2} \vec{f}_{\mu k}^T(\vec{\theta}_\mu) R^{-1} \vec{f}_{\mu k}(\vec{\theta}_\mu)].$$

В случае неинтерферирующих полей $f_{\mu'}$ и f_μ второй слагаемое исчезает и функция $g(\vec{u}_k, \vec{\theta})$ может быть записана в виде

$$g_0(\vec{u}_k, \vec{\theta}) = \sum_{\mu=1}^M \sum_k [\vec{u}_k^T R^{-1} \vec{f}_{\mu k}(\vec{\theta}_\mu) - \frac{1}{2} \vec{f}_{\mu k}^T(\vec{\theta}_\mu) R^{-1} \vec{f}_{\mu k}(\vec{\theta}_\mu)].$$

Представим полную функцию $g(\vec{u}_k, \vec{\theta})$ в записи аналогичной записи функции $g_0(\vec{u}_k, \vec{\theta})$. Мы выделим в функции $g(\vec{u}_k, \vec{\theta})$ член, содержащий число слоев m

$$\begin{aligned} g(\vec{u}_k, \vec{\theta}) = & \left[\sum_k \vec{u}_k^T R^{-1} \vec{f}_{mk}(\vec{\theta}_m) - \sum_{\substack{\mu=1 \\ \mu \neq m}}^M \sum_k \vec{f}_{\mu k}(\vec{\theta}_\mu) R^{-1} \vec{f}_{mk}(\vec{\theta}_m) - \right. \\ & \left. - \frac{1}{2} \sum_k \vec{f}_{mk}^T(\vec{\theta}_m) R^{-1} \vec{f}_{mk}(\vec{\theta}_m) \right] + \left[\sum_{\substack{\mu=1 \\ \mu \neq m}}^M \sum_k \vec{u}_k^T R^{-1} \vec{f}_{\mu k}(\vec{\theta}_\mu) - \right. \\ & \left. - \frac{1}{2} \sum_{\substack{\mu=1 \\ \mu \neq m}}^M \sum_{\substack{\mu'=1 \\ \mu' \neq m}}^M \sum_k \vec{f}_{\mu' k}^T(\vec{\theta}_{\mu'}) R^{-1} \vec{f}_{\mu k}^T(\vec{\theta}_\mu) - \frac{1}{2} \sum_{\substack{\mu=1 \\ \mu \neq m}}^M \sum_k \vec{f}_{\mu k}^T(\vec{\theta}_\mu) R^{-1} \vec{f}_{\mu k}(\vec{\theta}_\mu) \right]. \end{aligned}$$

Только выражение, заключенное в первые квадратные скобки содержит вектор параметров $\vec{\theta}_m$, поэтому для оценивания вектора $\vec{\theta}_m$ достаточно минимизировать функцию

$$\begin{aligned} g_m(\vec{u}_k, \vec{\theta}_m) = & \sum_k \vec{Y}_{km}^T R^{-1} \vec{f}_{mk}(\vec{\theta}_m) - \frac{1}{2} \sum_k \vec{f}_{mk}^T(\vec{\theta}_m) R^{-1} \vec{f}_{mk}(\vec{\theta}_m), \\ \vec{Y}_{km}^T = & \vec{u}_k^T - \sum_{\substack{\mu=1 \\ \mu \neq m}}^M \vec{f}_{k\mu}^T(\vec{\theta}_\mu) \end{aligned}$$

Такое представление является основой для организации итерационной процедуры.

Шаг 1. Предположим, что имеется только один сигнал:

$$\begin{aligned} M = 1, \quad g_1(\vec{u}_k, \vec{\theta}_1) = & \sum_k [\vec{u}_k^T R^{-1} \vec{f}_{1k}(\vec{\theta}_1) - \frac{1}{2} \vec{f}_{1k}^T(\vec{\theta}_1) R^{-1} \vec{f}_{1k}(\vec{\theta}_1)] \\ \hat{\vec{\theta}}_1^{(1)} = & \max_{\vec{\theta}_1} g_1(\vec{u}_k, \vec{\theta}_1). \end{aligned}$$

Проверка выполнения пороговых условий:

Если

$$\alpha_1^2 = \sum_k \vec{f}_{1k}^T(\hat{\vec{\theta}}_1^{(1)}) R^{-1} \vec{f}_{1k}(\hat{\vec{\theta}}_1^{(1)}) \geq \alpha_{0P_e}^2,$$

то выполнение процедуры продолжается.

Если

$$\alpha_1^2 < \alpha_{0P_e}^2,$$

то выполнение процедуры прекращается и можно говорить: отсутствуют какие-либо сигналы.

Шаг 2. Предположение о существовании двух сигналов:

$$M = 2, \quad g_2(\vec{u}_k, \hat{\theta}_1, \vec{\theta}_2) = \sum_k \vec{Y}_{k2}^T R^{-1} \vec{f}_{2k}(\vec{\theta}_2) - \frac{1}{2} \sum_k \vec{f}_{2k}^T(\vec{\theta}_2) R^{-1} \vec{f}_{2k}(\vec{\theta}_2)$$

$$\vec{Y}_{k2}^T = \vec{u}_k^T - \vec{f}_{1k}^T(\hat{\theta}_1).$$

Для вектора параметров $\vec{\theta}_1$ используется оценка, полученная на первом шаге $\hat{\theta}_1$

$$\hat{\theta}_2^{(1)} = \max_{\vec{\theta}_2} g_2(\vec{u}_k, \hat{\theta}_1, \vec{\theta}_2).$$

Проверка пороговых условий:

Если

$$\alpha_2^2 = \sum_k \vec{f}_{2k}^T(\hat{\theta}_2^{(1)}) R^{-1} \vec{f}_{2k}(\hat{\theta}_2^{(1)}) \geq \alpha_{0P_e}^2,$$

то итерационный процесс продолжается.

Если

$$\alpha_2^2 < \alpha_{0P_e}^2,$$

то выполнение процедуры прекращается и можно говорить: имеется только один сигнал $\hat{\theta}_1$.

Оцениваем, считая что сигналов два, оцениваем вектор параметров $\vec{\theta}_1$, отвечающий первому сигналу, фиксируя вектор параметров $\vec{\theta}_2$, отвечающий второму сигналу $\hat{\theta}_2^{(1)}$

$$g_2(\vec{u}_k, \vec{\theta}_1, \hat{\theta}_2^{(1)}) = \sum_k \vec{Y}_{k1}^T R^{-1} \vec{f}_{1k}(\vec{\theta}_1) - \frac{1}{2} \sum_k \vec{f}_{1k}^T(\vec{\theta}_1) R^{-1} \vec{f}_{1k}(\vec{\theta}_1),$$

$$\hat{\theta}_1^{(1)} = \max_{\vec{\theta}_1} g_2(\vec{u}_k, \vec{\theta}_1, \hat{\theta}_2^{(1)}), \quad \vec{Y}_{k1} = \vec{u}_k^T - \vec{f}_{2k}^T(\hat{\theta}_2).$$

Затем снова ищем $\hat{\theta}_2^{(2)}$, предполагая $\hat{\theta}_1^{(1)}$ равным только что уточненным значениям. Каждый раз переходя к новой итерации мы проверяем пороговое условие.

Если

$$\frac{g_2(\hat{\theta}_1^{(n)}, \hat{\theta}_2^{(n)}) - g_2(\hat{\theta}_1^{(n-1)}, \hat{\theta}_2^{(n-1)})}{g_2(\hat{\theta}_1^{(n-1)}, \hat{\theta}_2^{(n-1)})} < \delta \sim 10^{-2} \div 10^{-3},$$

переходим к выполнению шага $M = 3$.

Если

$$\frac{\Delta g_2^{(n)}}{\Delta g_2^{(n-1)}} > \delta,$$

продолжаем итерационный процесс для векторов $\hat{\theta}_2$ и $\hat{\theta}_1$.

Шаг 3. Предположим, что имеется три сигнала

$$M = 3 \quad g_3(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \vec{\theta}_3) = \sum_k \vec{Y}_{k3}^T R^{-1} \vec{f}_{k3}(\vec{\theta}_3) - \\ - \frac{1}{2} \sum_k \vec{f}_{3k}^T(\vec{\theta}_3) R^{-1} \vec{f}_{3k}(\vec{\theta}_3) \\ \vec{Y}_{k3}^T = \vec{u}_k^T - \vec{f}_{1k}^T(\hat{\theta}_1) - \vec{f}_{2k}^T(\hat{\theta}_2) \\ \hat{\theta}_3^{(1)} = \max_{\vec{\theta}_3} g_3(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \vec{\theta}_3).$$

Проверка выполнения порогового условия

Если

$$\alpha_3^2 = \sum_k \vec{f}_{3k}^T(\hat{\theta}_3^{(1)}) R^{-1} \vec{f}_{3k}(\hat{\theta}_3^{(1)}) \geq \alpha_{0P_e}^2,$$

то вычислительный процесс продолжается

Если

$$\alpha_3^2 < \alpha_{0P_e}^2,$$

то итерационный процесс прекращается и можно говорить: имеются сигналы с параметрами $\hat{\theta}_1, \hat{\theta}_2$.

После этого мы ищем поправки к оценкам $\hat{\theta}_2^{(1)}, \hat{\theta}_1^{(1)}$ и продолжаем уточнение, проверяя выполнение порогового условия на каждом шаге.

Если

$$\frac{g_3(\hat{\theta}_1^{(n)}, \hat{\theta}_2^{(n)}, \hat{\theta}_3^{(n)}) - g_3(\hat{\theta}_1^{(n-1)}, \hat{\theta}_2^{(n-1)}, \hat{\theta}_3^{(n-1)})}{g_3(\hat{\theta}_1^{(n-1)}, \hat{\theta}_2^{(n-1)}, \hat{\theta}_3^{(n-1)})} < \delta,$$

то переходим к шагу $m = 4$.

Если

$$\frac{\Delta g_3^{(n)}}{\Delta g_3^{(n-1)}} > \delta,$$

то продолжаем итерационную процедуру для векторов $\vec{\theta}_1, \vec{\theta}_2$ и $\vec{\theta}_3$.

Шаг m . Предполагаем, что имеется m сигналов.

$$M = m, \quad g_m(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \dots, \vec{\theta}_m) = \sum_k \vec{Y}_{km-1}^T R^{-1} \vec{f}_{mk}(\vec{\theta}_m) - \\ - \frac{1}{2} \sum_k \vec{f}_{mk}^T(\vec{\theta}_m) R^{-1} \vec{f}_{mk}(\vec{\theta}_m) \\ \vec{Y}_{km}^T = \vec{u}_k^T - \sum_{\mu=1}^{m-1} \vec{f}_{\mu k}^T(\hat{\theta}_\mu) \\ \hat{\theta}_m^{(1)} = \max_{\vec{\theta}_m} g_m(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \dots, \vec{\theta}_m).$$

Проверка порогового условия

Если

$$\alpha_m^2 = \sum_k \vec{f}_{mk}^T(\hat{\theta}_m^{(1)}) R^{-1} \vec{f}_{mk}(\hat{\theta}_m^{(1)}) \geq \alpha_{0Pe}^2,$$

то вычислительная процедура продолжается.

Если

$$\alpha_m^2 < \alpha_{0Pe}^2,$$

то процесс прекращается и можно говорить: имеется $(m - 1)$ сигналов с параметрами $\hat{\theta}_1, \dots, \hat{\theta}_{m-1}$.

После этого мы ищем поправки ко всем параметрам и используем для этого итерационную процедуру. Мы имеем $\hat{\theta}_1^{(1)}, \dots, \hat{\theta}_{m-1}^{(1)}, \hat{\theta}_m^{(1)}$. После следующей итерации получаем $\hat{\theta}_1^{(2)}, \dots, \hat{\theta}_{m-1}^{(2)}, \hat{\theta}_m^{(2)}$ и т.д. Каждый раз, переходя к следующей итерации, проверяется выполнение пороговых условий.

Если

$$\frac{g_m(\hat{\theta}_1^{(n)}, \dots, \hat{\theta}_{m-1}^{(n)}, \hat{\theta}_m^{(n)}) - g_m(\hat{\theta}_1^{(n-1)}, \dots, \hat{\theta}_{m-1}^{(n-1)}, \hat{\theta}_m^{(n-1)})}{g_m(\hat{\theta}_1^{(n-1)}, \dots, \hat{\theta}_{m-1}^{(n-1)}, \hat{\theta}_m^{(n-1)})} < \delta,$$

то переходим к шагу с $M = m + 1$.

Если

$$\frac{\Delta g_m^{(n)}}{\Delta g_m^{(n-1)}} > \delta,$$

то продолжаем итерационную процедуру для векторов $\vec{\theta}_1, \dots, \vec{\theta}_{m-1}, \vec{\theta}_m$.

Для приложений представляет большой интерес модификация этой итерационной процедуры без одновременной корректировки параметров. Рассмотрим модифицированную процедуру.

Шаг 1. Предполагаем, что имеется только один сигнал:

$$M = 1, \quad g_1(\vec{u}_k, \vec{\theta}_1) = \sum_k [\vec{u}_k^T R^{-1} \vec{f}_{1k}(\vec{\theta}_1) - \frac{1}{2} \vec{f}_{1k}^T(\vec{\theta}_1) R^{-1} \vec{f}_{1k}(\vec{\theta}_1)]$$

$$\hat{\theta}_1 = \max_{\theta_1} g_1(\vec{u}_k, \vec{\theta}_1).$$

Проверка порогового условия:

Если

$$\alpha_1^2 = \sum_k \vec{f}_k^T(\hat{\theta}_1) R^{-1} \vec{f}_k(\hat{\theta}_1) \geq \alpha_{0Pe}^2,$$

то продолжаем выполнение процедуры.

Если

$$\alpha_1^2 < \alpha_{0P_e}^2,$$

останавливаем выполнение процедуры и утверждаем: какие-либо сигналы отсутствуют.

Шаг 2. Предполагаем, что имеется два сигнала:

$$M = 2, \quad g_2(\vec{u}_k, \hat{\theta}_1, \vec{\theta}_2) = \sum_k \vec{Y}_{k2}^T R^{-1} \vec{f}_{2k}(\vec{\theta}_2) - \frac{1}{2} \sum_k \vec{f}_{2k}^T(\vec{\theta}_2) R^{-1} \vec{f}_{2k}(\vec{\theta}_2)$$

$$\vec{Y}_{k2}^T = \vec{u}_k^T - \vec{f}_{1k}^T(\hat{\theta}_1).$$

Для вектора параметров $\vec{\theta}_1$ используем оценку, полученную на первом шаге $\hat{\theta}_1$

$$\hat{\theta}_2 = \max_{\vec{\theta}_2} g_2(\vec{u}_k, \hat{\theta}_1, \vec{\theta}_2).$$

Проверка порогового условия:

Если

$$\alpha_2^2 = \sum_k \vec{f}_{2k}^T(\hat{\theta}_2) R^{-1} \vec{f}_{2k}(\hat{\theta}_2) \geq \alpha_{0P_e}^2,$$

то продолжаем выполнение процедуры.

Если

$$\alpha_2^2 < \alpha_{0P_e}^2,$$

то прекращаем выполнение процедуры и утверждаем: имеется только один сигнал $\vec{\theta}_1$.

Шаг 3. Предполагаем, что имеется три сигнала

$$M = 3, \quad g_3(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \vec{\theta}_3) = \sum_k \vec{Y}_{k3}^T R^{-1} \vec{f}_{3k}(\vec{\theta}_3) -$$

$$- \frac{1}{2} \sum_k \vec{f}_{3k}^T(\vec{\theta}_3) R^{-1} \vec{f}_{3k}(\vec{\theta}_3)$$

$$\vec{Y}_{k3}^T = \vec{u}_k^T - \vec{f}_{3k}^T(\hat{\theta}_1) - \vec{f}_{3k}^T(\hat{\theta}_2)$$

$$\hat{\theta}_3 = \max_{\vec{\theta}_3} g_3(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \vec{\theta}_3).$$

Проверка порогового условия.

Если

$$\alpha_3^2 = \sum_k \vec{f}_k^T(\hat{\theta}_3) R^{-1} \vec{f}_k(\hat{\theta}_3) \geq \alpha_{0P_e}^2,$$

то продолжаем выполнение процедуры.

Если

$$\alpha_3^2 < \alpha_{0P_e}^2,$$

то выполнение процедуры прекращается и мы утверждаем: имеются сигналы с параметрами $\hat{\theta}_1, \hat{\theta}_2$.

Шаг m . Предполагаем, что имеется m сигналов.

$$M = m, \quad g_m(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \dots, \vec{\theta}_m) = \sum_k \vec{Y}_{km-1}^T R^{-1} \vec{f}_{mk}(\vec{\theta}_m) - \\ - \frac{1}{2} \sum_k \vec{f}_{mk}^T(\vec{\theta}_m) R^{-1} \vec{f}_{mk}(\vec{\theta}_m) \\ \vec{Y}_{km-1}^T = \vec{u}_k^T - \sum_{\mu=1}^{m-1} \vec{f}_{\mu k}^T(\hat{\theta}_\mu) \\ \hat{\theta}_m = \max_{\vec{\theta}_m} g_m(\vec{u}_k, \hat{\theta}_1, \hat{\theta}_2, \dots, \vec{\theta}_m).$$

Проверка порогового условия.

Если

$$\alpha_m^2 = \sum_k \vec{f}_{mk}^T(\hat{\theta}_m) R^{-1} \vec{f}_{mk}(\hat{\theta}_m) \geq \alpha_{0Pe}^2,$$

то выполнение процедуры продолжается.

Если

$$\alpha_m^2 < \alpha_{0Pe}^2,$$

то выполнение процедуры прекращается и мы утверждаем: имеется $(m-1)$ с параметрами $\hat{\theta}_1, \dots, \hat{\theta}_{m-1}$.

Пример 5: В рассматриваемом случае, относящемся к сейсмике, запишем решающую функцию и оценки параметров для произвольного шага m модифицированной процедуры. Пусть случайная компонента некоррелирована ($R = \sigma^2 I$) а сигнал представляется выражением

$$\vec{f}_k(\vec{\theta}_m) = \|A_m \varphi(t_i - \tau_m - k \Delta x \gamma_m)\|_{i=1}^n,$$

$$M = m, \quad g_m(\vec{u}_k, A_m, \tau_m, \gamma_m) = \frac{A_m}{\sigma^2} \sum_i \varphi(t_i) \sum_k y(t_i + \tau_{mk}) - \\ - \frac{A_m^2}{2\sigma^2} K \sum_i \varphi^2(t_i),$$

где

$$y(t_i + \tau_{mk}) = u(t_i + \tau_{mk}) - \sum_{\mu' \neq m} A_{\mu'} \varphi(t_i - \tau_{\mu'k} - \tau_{\mu k}).$$

Можно найти явное выражение для оценки амплитуды

$$\frac{\partial g_m}{\partial A_m} = 0 \Rightarrow \hat{A}_m = \frac{\sum_i \varphi(t_i) \sum_k y(t_i + \tau_{mk})}{\sum_i \varphi^2(t_i)}.$$

Подставляя оценку \hat{A}_m в решающую функцию $g_m(\vec{u}_k, A_m, \tau_m, \gamma_m)$, получаем

$$g_m(\tau_m, \gamma_m) = \frac{1}{\sigma^2} \frac{[\sum_i \varphi(t_i) \sum_k y(t_i + \tau_{mk})]^2}{K \sum_i \varphi^2(t_i)}.$$

Оценки времени распространения τ_m и обратной кажущейся скорости γ_m для сигнала m можно записать следующим образом

$$\hat{\tau}_m, \hat{\gamma}_m = \max_{\tau_m, \gamma_m} g_m(\vec{u}_k, \hat{A}_m, \tau_m, \gamma_m).$$

Пороговое условие.

Если

$$\alpha_m^2 = \frac{\hat{A}_m^2}{\sigma^2} \sum_k \sum_i \varphi^2(t_i - \hat{\tau}_m - k\Delta x \hat{\gamma}_m) \geq \alpha_{0P_e}^2,$$

то переходим к следующему шагу $M = m + 1$.

Если

$$\alpha_m^2 < \alpha_{0P_e}^2,$$

то прекращаем выполнение процедуры.

Оценка вектора параметров имеет вид

$$\hat{\theta} \{ \hat{A}_1, \hat{\tau}_1, \hat{\gamma}_1, \hat{A}_2, \hat{\tau}_2, \hat{\gamma}_2, \dots, \hat{A}_m, \hat{\tau}_m, \hat{\gamma}_m \}.$$

Пример 6: Аналогично, в случае магнитного поля, запишем решающую функцию и оценку параметров для произвольного шага m (число сигналов – μ) при условии некоррелированности случайной компоненты $\varepsilon - R = \sigma^2 I$

$$f_{k\mu} = \frac{I_\mu h_\mu}{h_\mu^2 + (x_k - \xi_\mu^2)} = I_\mu \varphi_k(h_\mu, \xi_\mu),$$

где

$$\varphi_k(h_\mu, \xi_\mu) = \frac{h_\mu}{h_\mu^2 + (x_k - \xi_\mu^2)}.$$

$$M = m, \quad g_m(u_k, I_m, h_m, \xi_m) = \frac{I_m}{\sigma^2} \sum_k y_{km} \varphi(h_m, \xi_m) - \frac{1}{2} \frac{I_m^2}{\sigma^2} \sum_k \varphi_k^2(h_m, \xi_m),$$

где

$$y_{km} = u_k - \sum_{\mu' \neq m} I_{\mu'} \varphi_k(h_{\mu'}, \xi_{\mu'}).$$

Явное выражение для оценки намагниченности:

$$\frac{\partial g_m}{\partial I_m} = 0, \quad \hat{I}_m = \frac{\sum_k y_{km} \varphi_k(h_m, \xi_m)}{\sum_k \varphi_k^2(h_m, \xi_m)}.$$

Подставляя оценку \hat{I}_m в решающую функцию $g_m(u_k, I_m, h_m, \xi_m)$ получаем

$$g_m(h_m, \xi_m) = \frac{1}{\sigma^2} \frac{\sum_k y_{km} \varphi_k(h_m, \xi_m)}{\sum_k \varphi_k^2(h_m, \xi_m)}.$$

Оцениваемые параметры h_m (глубина верхней грани) и ξ_m (горизонтальная координата) m вертикальных магнитных слоев находим максимизируя функцию g_m

$$h_m, \xi_m = \max g_m(u_k, \hat{I}_m, h_m, \xi_m).$$

Пороговое условие записывается в виде

Если

$$\alpha_m^2 = \frac{\hat{I}_m^2}{\sigma_m^2} \sum_k \varphi_k^2(\hat{h}_m, \hat{\xi}_m) \geq \alpha_{0P_e}^2,$$

то переходим к следующему шагу $M = m + 1$.

Если

$$\alpha_m^2 < \alpha_{0P_e}^2,$$

выполнение процедуры прекращается.

Полученные оценки вектора параметров имеют вид

$$\hat{\theta} = \{\hat{I}_1, \hat{h}_1, \hat{\xi}_1, \hat{I}_2, \hat{h}_2, \hat{\xi}_2, \dots, \hat{I}_m, \hat{h}_m, \hat{\xi}_m\}.$$

Глава 9

Геофизическая лучевая томография

9.1 Основные определения

Пусть $f(\vec{r})$ – неизвестная функция, а величина u_i – ее интегральная проекция для пары точек с номером i , которые будем называть источником (source) и приемником (receiver) $(s, r)_i$. Связь между u_i и $f(\vec{r})$ – интеграл вдоль луча L_i

$$u_i = \int_{L_i} f(\vec{r}) dl.$$

Последний интеграл можно записать в виде

$$u_i = \int_{\Omega} f(\vec{r}) \delta(L_i(\vec{r})) d\vec{r},$$

где Ω – область интегрирования (поверхность или объем), $\delta(\cdot)$ δ -функция. Обозначим интегральный оператор как \tilde{R}

$$u_i = \tilde{R}_i[f(\vec{r})], \quad \tilde{R}_i[\cdot] \equiv \int d\vec{r} \delta(L_i(\vec{r}))[\cdot].$$

После введения полярных координат, в двумерном случае неизвестную функцию $f(\vec{r})$ можно записать как $f(r, \varphi)$, используя известное соотношение $\vec{r} = (r, \varphi)$. Совершив переход к паре параметров (n, θ) получим прямое преобразование Радона

$$(n, \theta), \quad u(n, \theta) = \int_L f(\sqrt{l^2 + n^2}, \theta - \arctan(l/n)) dl.$$

Параметры, связанные с типичными для лучевой томографии преобразованиями, представлены на рис. 9.1.

9.2 Преобразование Радона

Прямое преобразование Радона определяется следующим образом

$$u(n, \theta) = \tilde{R}[f(r, \varphi)].$$

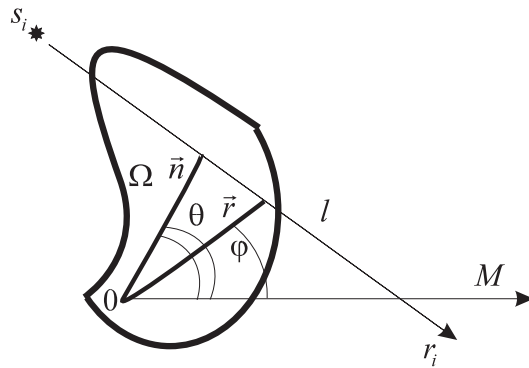


Рис. 9.1: Основные параметры, характерные для преобразований, встречающихся в лучевой томографии.

Обратное преобразование Радона имеет вид

$$f(r, \varphi) = \frac{1}{2\pi^2} \int_{-\infty}^{\infty} \int_0^{\pi} \frac{1}{r \cos(\theta - \varphi) - n} \frac{\partial u(n, \theta)}{\partial n} dn d\theta$$

или, в операторной форме:

$$f(r, \varphi) = \tilde{R}^{-1}[u(n, \theta)].$$

Оператор обратного преобразования Радона может быть представлен как последовательность операторов:

$$\tilde{R}^{-1} = -\frac{1}{2\pi} PHD,$$

где

$$D : q(n, \theta) = D[u(n, \theta)] = \lim_{\Delta n \rightarrow 0} \frac{u(n + \Delta n, \theta) - u(n, \theta)}{\Delta n}$$

– дифференциальный оператор (первая производная);

$$H : h(n', \theta) = H[q(n', \theta)] = -\frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \left\{ \int_{-\infty}^{n' - \varepsilon} \frac{q(n, \theta)}{n' - n} dn + \int_{n' + \varepsilon}^{\infty} \frac{q(n, \theta)}{n' - n} dn \right\}$$

– преобразование Гильберта;

$$P : p(r, \varphi) = P[h(n', \theta)] = \int_0^{\pi} h(r \cos(\theta - \varphi), \theta) d\theta$$

– оператор обратного проецирования.

Преобразование Радона имеет ограниченную область применимости, поскольку на практике мы никогда не встречаем представление данных в непрерывном виде. Рассмотрим алгебраические методы, имеющие существенно большее значение при решении практических задач.

9.3 Алгебраические методы

Представим неизвестную функцию $f(\vec{r})$ в виде разложения по базисным функциям

$$f(\vec{r}) = \sum_{j=1}^M \theta_j \psi_j(\vec{r}).$$

Данные наблюдений u_i могут быть записаны в этом случае следующим образом

$$u_i = \sum_{j=1}^M C_{ij} \theta_j,$$

где

$$C_{ij} = \tilde{R}_i[\psi_j(\vec{r})].$$

Простейшие базисные функции:

$$\psi_j(\vec{r}) = \begin{cases} 1 & \text{для } \vec{r} \in \Omega_j, \\ 0 & \text{для } \vec{r} \notin \Omega_j. \end{cases}$$

Элементы C_{ij} в этом случае записываются следующим образом

$$C_{ij} = \begin{cases} C & \text{для } L_i \cap \Omega_j \neq \emptyset, \\ 0 & \text{для } L_i \cap \Omega_j = \emptyset. \end{cases}$$

Основные элементы схемы реализации алгебраических методов представлены на рис. 9.2.

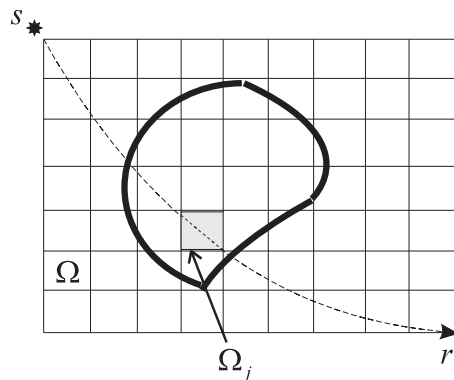


Рис. 9.2: Основные элементы схемы реализации алгебраических методов.

Запишем модель измеряемых величин в виде

$$\vec{u} = C\vec{\theta} + \vec{\varepsilon},$$

где $\vec{\varepsilon}$ – случайная компонента ($\vec{\varepsilon} \in N(0, R_\varepsilon)$) и $\vec{\theta}$ – случайный вектор параметров ($\vec{\theta} \in N(\langle\vec{\theta}\rangle, R_\theta)$).

Задача состоит в получении оценки для вектора параметров $\vec{\theta}$. Для этого будем использовать метод максимума апостериорной вероятности:

$$\tilde{\vec{\theta}} = \max_{\vec{\theta}} p(\vec{\theta})p(\vec{u}/\vec{\theta}).$$

В случае нормального распределения векторов $\vec{\varepsilon}$ и $\vec{\theta}$ такой критерий принимает вид

$$\tilde{\vec{\theta}} = \min_{\vec{\theta}} [(\vec{u} - C\vec{\theta})^T R_\varepsilon^{-1} (\vec{u} - C\vec{\theta}) + (\vec{\theta} - \langle\vec{\theta}\rangle)^T R_\theta^{-1} (\vec{\theta} - \langle\vec{\theta}\rangle)].$$

Явная запись искомой оценки и ее ковариационная матрица могут быть представлены в двух видах:

1. $\hat{\vec{\theta}} = (C^T R_\varepsilon^{-1} C + R_\theta^{-1})^{-1} (C^T R_\varepsilon^{-1} \vec{u} + R_\theta^{-1} \langle\vec{\theta}\rangle),$
 $R(\hat{\vec{\theta}}) = (C^T R_\varepsilon^{-1} C + R_\theta^{-1})^{-1}.$
2. $\tilde{\vec{\theta}} = \langle\vec{\theta}\rangle + R_\theta C^T (C R_\theta C^T + R_\varepsilon)^{-1} (\vec{u} - C\langle\vec{\theta}\rangle),$
 $R(\tilde{\vec{\theta}}) = R_\theta - R_\theta C^T (C R_\theta C^T + R_\varepsilon)^{-1} C R_\theta.$

В частном случае некоррелированных векторов $R_\varepsilon = \sigma_\varepsilon^2 I$ и $R_\theta = \sigma_\theta^2 I$, оценка и соответствующая ей ковариационная матрица могут быть представлены в виде

1. $\hat{\vec{\theta}} = (C^T C + \alpha I)^{-1} (C^T \vec{u} + \alpha \langle\vec{\theta}\rangle),$
 $R(\hat{\vec{\theta}}) = \sigma_\varepsilon^2 (C^T C + \alpha I)^{-1},$

где $\alpha = \sigma_\varepsilon^2 / \sigma_\theta^2$ регуляризирующий параметр (по Тихонову).

2. $\tilde{\vec{\theta}} = \langle\vec{\theta}\rangle + C^T (C C^T + \alpha I)^{-1} (\vec{u} - C\langle\vec{\theta}\rangle),$
 $R(\tilde{\vec{\theta}}) = \sigma_\theta^2 - \sigma_\theta^2 C^T (C C^T + \alpha I)^{-1} C.$

Ковариационные матрицы $R(\hat{\theta})$ и $R(\tilde{\theta})$ могут использоваться для оценки качества интерпретации. Дисперсии ошибок восстановления расположены на главных диагоналях матриц $R(\hat{\theta})$ и $R(\tilde{\theta})$, при этом недиагональные члены описывают корреляционные связи между элементами оцениваемых векторов.

Рассмотрим рекуррентный алгоритм, который учитывает разреженный характер матрицы C .

В качестве начального вектора параметров $\vec{\theta}^{(0)}$ и начальной ковариационной матрицы $R_\theta^{(0)}$ мы возьмем априорный вектор

$$\vec{\theta}^{(0)} = \langle \vec{\theta} \rangle$$

и априорную ковариационную матрицу

$$R_\theta^{(0)} = R_\theta.$$

Первое приближение для оценки $\vec{\theta}$ вычисляется по формуле

$$\vec{\theta}^{(1)} = \vec{\theta}^{(0)} + \alpha^{(1)} \Delta \vec{\theta}^{(1)},$$

где

$$\begin{aligned} \Delta \vec{\theta}^{(1)} &= R_\theta^{(0)} C^{(1)T}, & \alpha^{(1)} &= e_1 / \delta_1^2, \\ e_1 &= u_1 - C^{(1)} \vec{\theta}^{(1)}, & \delta_1^2 &= \sigma_{\varepsilon_1}^2 + C^{(1)} \Delta \vec{\theta}^{(1)}, \end{aligned}$$

где $C^{(1)}$ – первая строка матрицы C . Первое приближение для ковариационной матрицы дается выражением

$$R_\theta^{(1)} = R_\theta^{(0)} - \Delta \vec{\theta}^{(1)} \Delta \vec{\theta}^{(1)T} / \delta_1^2.$$

Пусть имеется $(i - 1)$ приближение

$$\vec{\theta}^{(i-1)}, \quad R_\theta^{(i-1)}.$$

Получим i -е приближение оценки $\vec{\theta}$

$$\vec{\theta}^{(i)} = \vec{\theta}^{(i-1)} + \alpha^{(i)} \Delta \vec{\theta}^{(i)},$$

где

$$\alpha^{(i)} = e_i / \delta_i^2, \quad e_i = u_i - C^{(i)} \vec{\theta}^{(i-1)}, \quad \delta_i^2 = \sigma_{\varepsilon_i}^2 + C^{(i)} \Delta \vec{\theta}^{(i)}, \quad \Delta \vec{\theta}^{(i)} = R_\theta^{(i-1)} C^{(i)T},$$

$C^{(i)}$ – i -я строка матрицы C . Далее, i -е приближение для ковариационной матрицы записывается в виде

$$R_\theta^{(i)} = R_\theta^{(i-1)} - \Delta \vec{\theta}^{(i)} \Delta \vec{\theta}^{(i)T} / \delta_i^2.$$

Число аппроксимаций не равно числу лучей.

Можно оценить качество аппроксимации по отношению к априорным представлениям на каждом i -м шаге, используя информацию относительно параметра θ_j

$$I_j^{(i)} = \frac{1}{2} \ln \frac{R_{\theta jj}^{(0)}}{R_{\theta jj}^{(i)}}, \quad j = 1, 2, \dots, M.$$

Влияние введения нового луча на величину информации описывается формулой

$$I_j^{(i-1) \rightarrow (i)} = \frac{1}{2} \ln \frac{R_{\theta jj}^{(i-1)}}{R_{\theta jj}^{(i)}}, \quad j = 1, 2, \dots, M.$$

9.4 Восстановление медленности и коэффициента поглощения

Пусть источники расположены на плоскости $z = 0$, а геофоны – в скважине. Предположим, что имеется априорная информация относительно сейсмической среды в форме медленности $S_0(r) = 1/v_0(r)$ и коэффициента поглощения $\alpha_0(r)$. Схема применения лучевой томографии представлена на рис. 9.3.

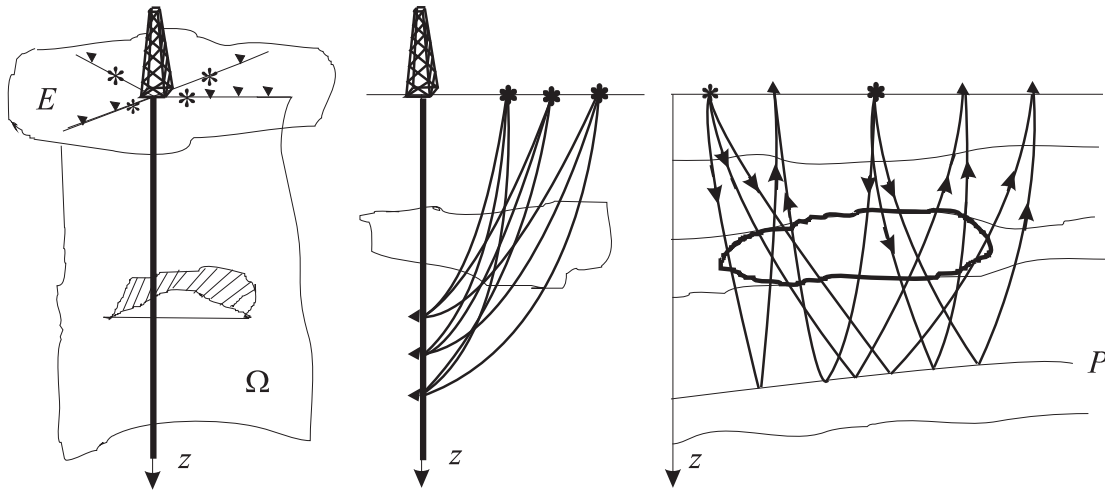


Рис. 9.3: Схема применения лучевой томографии.

Можно вычислить спектр сейсмического поля для каждой пары источник-приемник в сейсмической среде с функциями $S_0(r)$ и $\alpha_0(r)$:

$$u_{0i}(\omega) = A_i(\omega, r_{ri}, r_{si}, \alpha_0(r), S_0(r)) \exp(-i\omega\tau_0(r_{ri}, r_{si})),$$

где $(r_{ri}$ и $r_{si})$ – координаты соответственно приемника и источника, τ_0 – время распространения волны. Отметим, что функционал A_i описывает решение прямой задачи. Предположим, что отклонения $\delta\alpha$ и δS вещественных функций $\alpha(r)$

и $S(r)$ от $\alpha_0(r)$ и $S_0(r)$ малы и можно применять теорию возмущений

$$\begin{aligned}\delta\alpha(r) &= \alpha(r) - \alpha_0(r), & \delta\alpha(r) &\ll \alpha_0(r). \\ \delta S(r) &= S(r) - S_0(r), & \delta S(r) &\ll S_0(r).\end{aligned}$$

Умножим функцию на

$$\exp(-i\omega\tau_1),$$

комплексная величина $\tau_1 = \tau_{1\Re} + i\tau_{1\Im}$ с вещественной частью $\tau_{1\Re}$ определяет поправку на время распространения волны, а мнимая часть $\tau_{1\Im}$ определяет поправку к коэффициенту затухания. Пусть полное время

$$\tau = \tau_0 + \tau_1$$

удовлетворяет уравнению Эйконала

$$(\vec{\nabla}\tau)^2 = S^2(r),$$

при этом $S(r) = S_0(r) + S_1(r)$ и $S_1(r) = S_{1\Re} + S_{1\Im}$, где $S_{1\Re}$ и $S_{1\Im}$ соответственно вещественная и мнимая части комплексной функции $S_1(r)$. Теперь можно записать

$$(\vec{\nabla}(\tau_0 + \tau_1))^2 = (S_0 + S_1)^2$$

и

$$(\vec{\nabla}\tau_0)^2 + 2\vec{\nabla}\tau_0\vec{\nabla}\tau_1 + (\vec{\nabla}\tau_1)^2 = S_0^2 + 2S_0S_1 + S_1^2.$$

С учетом равенств

$$(\vec{\nabla}\tau_1)^2 \approx 0, \quad S_1^2 \approx 0, \quad (\vec{\nabla}\tau_0)^2 = S_0^2,$$

получим

$$\vec{\nabla}\tau_0\vec{\nabla}\tau_1 = S_0S_1.$$

Воспользовавшись равенством

$$\vec{\nabla}\tau_0 = S_0\vec{l}$$

где \vec{l} – единичный вектор касательный к лучу, получим

$$\vec{\nabla}_l\tau_1 = S_1 \Rightarrow \vec{\nabla}_l\tau_{1\Re} = S_{1\Re}, \quad \vec{\nabla}_l\tau_{1\Im} = S_{1\Im},$$

где $\vec{\nabla}_l$ – производная в направлении \vec{l} .

Интегрируя уравнения вдоль луча L получаем формулы

$$\tau_{1\Re} = \int_L S_{1\Re} dl, \quad \tau_{1\Im} = \int_L S_{1\Im} dl.$$

Запишем модель для времени распространения в виде

$$\tau_i - \tau_{i0} = \int_{L_i} S_{1\Re}(r) dl + \varepsilon_i.$$

Модель для наблюдаемого спектра амплитуды:

$$|u_i(\omega)|_{\max} = |u_{0i}(\omega)|_{\max} \exp(-\alpha_i(l)L_i) + \varepsilon_i,$$

где

$$\alpha_i(l) = \frac{1}{L_i} \int_{L_i} \alpha_l dl, \quad \varepsilon_i \ll |u_0(\omega)|_{\max} \exp(-\alpha_i L_i),$$

$$\ln \frac{|u_i(\omega)|_{\max}}{|u_{i0}(\omega)|_{\max}} = - \int_{L_i} \alpha_l(r) dl + \tilde{\varepsilon}_i,$$

где

$$\tilde{\varepsilon}_i = \frac{\varepsilon_i}{|u_{0i}(\omega)|_{\max} \exp(-\alpha_i L_i)}, \quad \ln(1 + \tilde{\varepsilon}_i) \approx \tilde{\varepsilon}_i.$$

Мы можем использовать эти модели для восстановления медленности $S(r)$ и коэффициента поглощения $\alpha_l(r)$.

9.5 Преобразование Радона сейсмограмм

Прямое преобразование сейсмограмм записывается в виде

$$s(p, \tau) = \int_{-\infty}^{\infty} u(x, px + \tau) dx$$

(см. рис. 9.4). Обратное преобразование Радона определяется как

$$u(x, t) = - \int_{-\infty}^{\infty} H \frac{d}{d\tau} s(p, \tau) dp = -PHD[s(p, \tau)]$$

или

$$u(x, t) = - \int_{-\infty}^{\infty} \left[-\frac{1}{\pi\tau} * \frac{\partial}{\partial\tau} s(p, \tau) \right] dp.$$

Запишем двумерное преобразование Фурье от сейсмограммы

$$u(k_x, \omega) = \iint_{-\infty}^{\infty} u(x, t) \exp[i(\omega t - k_x x)] dx dt$$

(см. рис. 9.5). Используя формулу $k_x = p\omega$ можно переписать выражение для спектра $u(k_x, \omega)$ в виде

$$u(p\omega, \omega) = \iint_{-\infty}^{\infty} u(x, t) \exp[i(\omega t - p\omega x)] dx dt$$

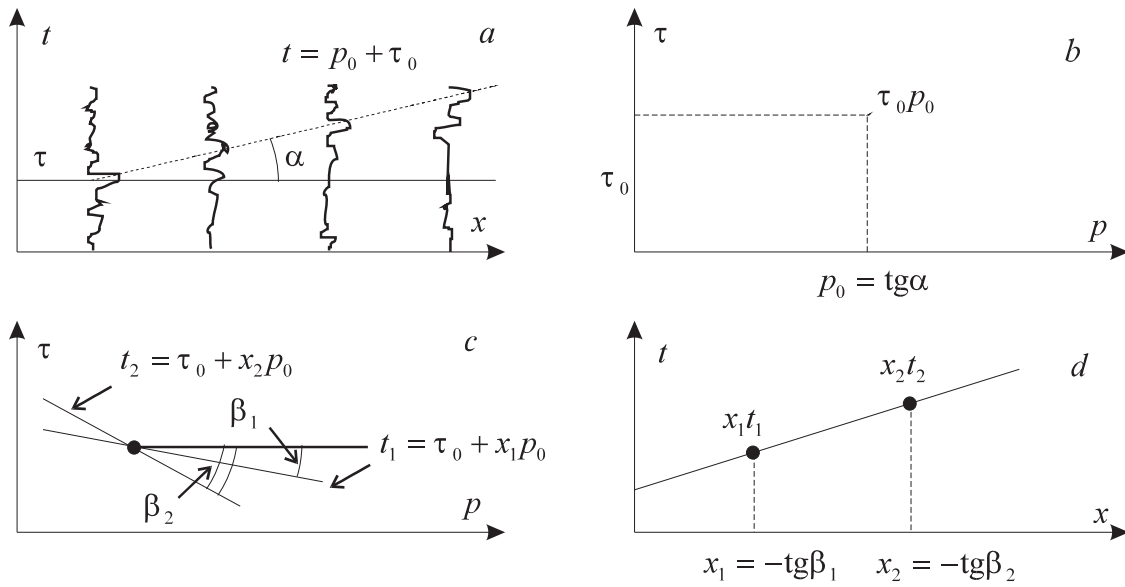


Рис. 9.4: Прямое и обратное преобразования Радона. Прямая линия в пространственно-временных координатах (а) и точка на плоскости с координатами угловой коэффициент-время отражения. Точка на плоскости с координатами угловой коэффициент-время отражения (с) и линия в пространственно-временных координатах (d).

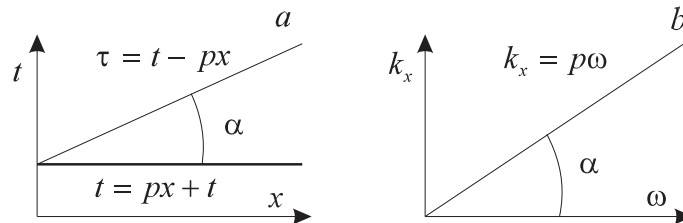


Рис. 9.5: Линия в пространственно-временных координатах (а) и линия на плоскости волновое число-частота.

или, вводя новую переменную $t = \tau + px$, получаем следующее выражение:

$$\begin{aligned}
 u(p\omega, \omega) &= \int_{-\infty}^{\infty} e^{i\omega\tau} \left[\int_{-\infty}^{\infty} u(x, \tau + px) dx \right] d\tau = \\
 &= \int_{-\infty}^{\infty} e^{i\omega\tau} s(p, \tau) d\tau = FR[u(x, t)],
 \end{aligned}$$

где $s(p, \tau) = \int_{-\infty}^{\infty} u(x, \tau + px) dx$ – преобразование Радона сейсмограммы u . Таким образом преобразование Радона и одномерное преобразование Фурье могут быть представлены двумерным преобразованием Фурье. Фильтрация в пространстве $(k - \omega)$ трансформируется в интегрирование в пределах от k_{min} до k_{max} (см.

рис. 9.6)

$$u(x) = \frac{1}{2\pi} \int_{k_{\min}}^{k_{\max}} u(k) \exp(ikx) dk,$$

где $p = |\omega|/p$ или

$$u(x) = \frac{|\omega|}{2\pi} \int_{p_{\min}}^{p_{\max}} u(p) \exp(i\omega p x) dp \quad (dk = |\omega| dp),$$

Фильтрацию сейсмограмм можно представить в виде

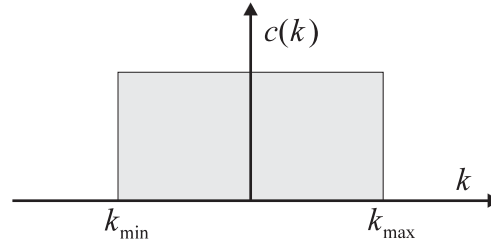


Рис. 9.6: Функция окна для интегрирования в k -области.

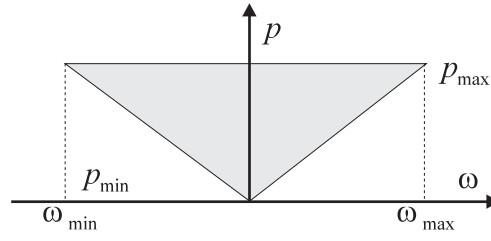
$$\begin{aligned} u(x, t) &= \frac{1}{4\pi^2} \int_{\omega_{\min}}^{\omega_{\max}} \int_{p_{\min}}^{p_{\max}} u(\omega, p) |\omega| \exp[-i\omega(t - px)] dp d\omega = \\ &= \frac{1}{2\pi} \int_{p_{\min}}^{p_{\max}} dp \left[\frac{1}{2\pi} \int_{\omega_{\min}}^{\omega_{\max}} \exp[-i\omega(t - px)] \times \right. \\ &\quad \left. \times FR[u(x, t)] |\omega| d\omega \right] = \\ &= \frac{1}{2\pi} \int_{p_{\min}}^{p_{\max}} R[u(x, t)] * (F^{-1}[|\omega|]) dp \end{aligned}$$

(см. рис. 9.7).

9.6 Преобразование Гильберта и аналитический сигнал

Предполагаем, что функция $S(t)$ имеет норму L^p , т.е.

$$S(t) \in L^p(-\infty, \infty), \quad \int_{-\infty}^{\infty} |S(t)|^p dt < \infty, \quad p \geq 1.$$

Рис. 9.7: Область интегрирования в плоскости (ω, k) .

В этом случае прямое и обратное преобразования Гильберта можно представить в виде

$$Q(t) = -\frac{1}{\pi} \text{V.P.} \int_{-\infty}^{\infty} \frac{S(\tau)}{t - \tau} d\tau = -S(t) * \frac{1}{\pi t}$$

и

$$S(t) = \frac{1}{\pi} \text{V.P.} \int_{-\infty}^{\infty} \frac{Q(\tau)}{t - \tau} d\tau,$$

где *V.P.*, что интегрирование в окрестности $t = \tau$ осуществляется в смысле главного значения. Найдем спектр функции $Q(t)$

$$Q(\omega) = \int_{-\infty}^{\infty} Q(t) \exp(i\omega t) dt = -\frac{1}{\pi} \int_{-\infty}^{\infty} S(\tau) \int_{-\infty}^{\infty} \frac{\exp(i\omega t)}{\tau - t} dt d\tau.$$

Сделав замену переменной $\tau - t = u$, получаем

$$\begin{aligned} Q(\omega) &= -\frac{1}{\pi} \int_{-\infty}^{\infty} S(\tau) \exp(i\omega\tau) \int_{-\infty}^{\infty} \frac{\exp(i\omega u)}{u} du d\tau = \\ &= -S(\omega) \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\exp(i\omega u)}{u} du = \\ &= -S(\omega) \frac{2i}{\pi} \int_0^{\infty} \frac{\sin \omega u}{u} du = -iS(\omega) \text{sign } \omega, \end{aligned}$$

т.к.

$$\int_0^{\infty} \frac{\sin \omega u}{u} du = \frac{\pi}{2} \text{sign } \omega$$

(см. рис. 9.8). Отметим, что

$$|Q(\omega)| = |S(\omega)|, \quad \arg Q(\omega) = \arg S(\omega) \pm \pi/2.$$

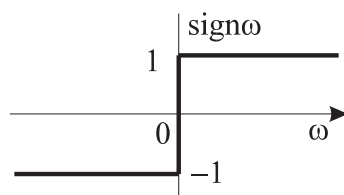


Рис. 9.8: Графическое представление функции $\text{sign } \omega$.

Величина $Q(\omega)$ является квадратурным фильтром, который изменяет фазу на $\pm\pi/2$ и не оказывает воздействия на амплитуду спектра.

Пусть $Z(t)$ – комплексная функция, зависящая от времени, – или аналитический сигнал

$$Z(t) = S(t) + iQ(t) \quad \text{или} \quad Z(t) = a(t) \exp(i\varphi(t)),$$

где

$$S(t) = \Re Z(t) = a(t) \cos \varphi(t),$$

$$Q(t) = \Im Z(t) = a(t) \sin \varphi(t).$$

Амплитуда и фаза записываются соответственно в виде

$$a(t) = (S^2(t) + Q^2(t))^{1/2} \quad \text{и} \quad \varphi(t) = \arctan(Q(t)/S(t)).$$

Амплитудная и фазовая функции аналитического сигнала находят применение используются при интерпретации сейсмических сигналов.

Глава 10

Дифракционная томография

Дифракционная томография является методом восстановления параметров среды с высоким пространственным разрешением. В дифракционной томографии для восстановления целевой функции используется рассеянное поле (дифрагированное поле).

10.1 Дифракционная томография в случае скалярного волнового уравнения

Для демонстрации основных положений дифракционной томографии рассмотрим случай акустического волнового уравнения в области свободной от источников.

$$\Delta u(\vec{x}, t) - \frac{1}{v^2(\vec{x})} \frac{\partial^2}{\partial t^2} u(\vec{x}, t) = 0, \quad x \in R^3,$$

где $u(\vec{x}, t)$ – поле давления, $v(\vec{x})$ – неизвестная скорость распространения акустической волны. Будем предполагать, что в опорной среде содержащей искомую неоднородность скорость распространения акустической волны является постоянной $v(\vec{x}) = const$ (см. рис. 10.1).

Вычислим спектр волны давления

$$u(\vec{x}, \omega) = \int_{-\infty}^{\infty} u(\vec{x}, t) \exp(i\omega t) dt.$$

Рассмотрим применение *метода возмущения* к скалярному волновому уравнению. Запишем обратное преобразование Фурье для поля давления

$$u(\vec{x}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(\vec{x}, \omega) \exp(-i\omega t) d\omega.$$

Подставляя выражение для $u(\vec{x}, t)$ в скалярное волновое уравнение, получаем уравнение Гельмгольца

$$\Delta u(\vec{x}, \omega) + \frac{\omega^2}{v^2(\vec{x})} u(\vec{x}, \omega) = 0.$$

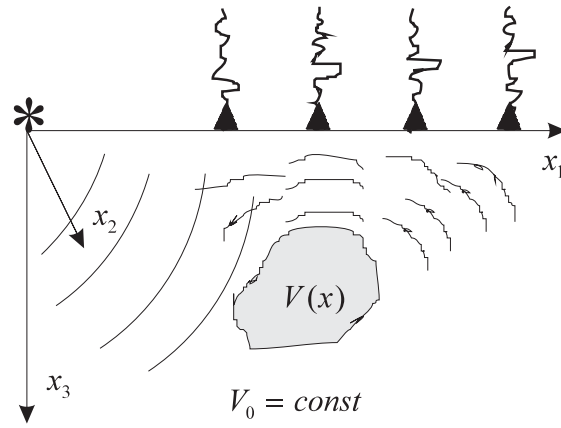


Рис. 10.1: Схема применения метода дифракционной томографии.

Добавляя член $(\omega^2/v^2(\vec{x}))u(\vec{x}, \omega)$ к левой и правой частям уравнения приходим к выражению

$$\Delta u(\vec{x}, \omega) + \frac{\omega^2}{v_0^2} u(\vec{x}, \omega) = \beta(\vec{x}, \omega) u(\vec{x}, \omega),$$

где

$$\beta(\vec{x}, \omega) = \frac{\omega^2}{v_0^2} \left(1 - \frac{v_0^2}{v^2(\vec{x})} \right) = \frac{\omega^2}{v_0^2} \beta_0(\vec{x}).$$

Выражение $\beta(\vec{x}, \omega)u(\vec{x}, \omega)$ можно рассматривать как некоторый источник. Будем искать решение в виде разложения, характерного для метода возмущений,

$$u(\vec{x}, \omega) = u_0(1 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots),$$

где ε – малый параметр ($\varepsilon \ll 1$). Подставляя разложение $u(\vec{x}, \omega)$ в уравнение Гельмгольца, получаем

$$\begin{aligned} \left(\Delta + \frac{\omega^2}{v_0^2} \right) u_0(1 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots) = \\ = \varepsilon \tilde{\beta} u_0(1 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots), \end{aligned}$$

где $\tilde{\beta} = \beta/\varepsilon \sim O(1)$.

Приравнивая члены с одинаковыми степенями ε

$$\begin{aligned} \varepsilon^0 : \quad \left(\Delta + \frac{\omega^2}{v_0^2} \right) u_0(\vec{x}, \omega) &= 0, \\ \varepsilon^1 : \quad \left(\Delta + \frac{\omega^2}{v_0^2} \right) u_0(\vec{x}, \omega) u_1(\vec{x}, \omega) &= \tilde{\beta} u_0(\vec{x}, \omega), \end{aligned}$$

и принимая во внимание

$$\begin{aligned} \Delta(u_0 u_1) &= u_1 \Delta u_0 + u_0 \Delta u_1 + 2(\vec{\nabla} u_0, \vec{\nabla} u_1), \\ \underbrace{u_1 \left(\Delta + \frac{\omega^2}{v^2} \right) u_0 + u_0 \Delta u_1}_{=0} + 2(\vec{\nabla} u_0, \vec{\nabla} u_1) &= \tilde{\beta} u_0, \end{aligned}$$

получаем

$$\varepsilon^1 : \quad \Delta u_1 + \frac{2(\vec{\nabla}u_0, \vec{\nabla}u_1)}{u_0} = \tilde{\beta}.$$

Аналогично легко выписываются рекурсионные соотношения для произвольного номера приближения

$$\begin{aligned} \Delta u_0 + \frac{\omega^2}{v_0^2} u_0 &= 0, \\ \Delta u_1 + \frac{2(\vec{\nabla}u_0, \vec{\nabla}u_1)}{u_0} &= \tilde{\beta}, \\ \Delta u_2 + \frac{2(\vec{\nabla}u_0, \vec{\nabla}u_2)}{u_0} &= \tilde{\beta}u_1, \\ \dots\dots\dots \\ \Delta u_n + \frac{2(\vec{\nabla}u_0, \vec{\nabla}u_n)}{u_0} &= \tilde{\beta}u_{n-1}. \end{aligned}$$

Рассмотрим падение плоской волны u_0 (см. рис. 10.2) на объект, расположенный в однородном безграничном пространстве:

$$u_0 = a_0 \exp[i(\vec{k}_0, \vec{x})], \quad \vec{k}_0 = \vec{n}\omega/v_0, \quad \vec{n} = \vec{\nabla}u/|\vec{\nabla}u|.$$

Ограничимся рассмотрением приближением первого порядка:

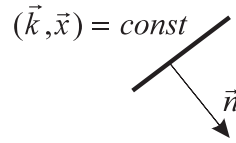


Рис. 10.2: Графическое представление соотношений между векторами \vec{k}_0 , \vec{x} и \vec{n} .

$$\Delta u_1 + 2i(\vec{k}_0, \vec{\nabla})u_1 = \tilde{\beta}.$$

Уравнение для функции Грина:

$$(\Delta + 2i(\vec{k}_0, \vec{\nabla}))G(\vec{x}, \vec{x}') = \delta(\vec{x} - \vec{x}').$$

Подставим выражение полученные обратным преобразование Фурье для функции Грина и δ -функции

$$\begin{aligned} G(\vec{x}, \vec{x}') &= \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} G(\vec{k}, \vec{x}') \exp[-i(\vec{k}, \vec{x})] d\vec{k}, \\ \delta(\vec{x} - \vec{x}') &= \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \exp[-i(\vec{k}, \vec{x} - \vec{x}')] d\vec{k}, \end{aligned}$$

в уравнение для функции Грина. После подстановки получим для спектра

$$(-k^2 + 2(\vec{k}_0, \vec{k}))G(\vec{k}, \vec{x}') = \exp[i(\vec{k}, \vec{x}')]]$$

и

$$G(\vec{k}, \vec{x}') = \frac{\exp[i(\vec{k}, \vec{x}')]]}{-k^2 + 2(\vec{k}_0, \vec{k})}.$$

Выполнив обратное преобразование Фурье, получим выражение для функции Грина в первом приближении

$$G(\vec{x}, \vec{x}') = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \frac{\exp\{i[(\vec{k}, \vec{x}') - (\vec{k}, \vec{x})]\}}{-k^2 + 2(\vec{k}_0, \vec{k})} d\vec{k}.$$

После замены переменных $\vec{k}' = \vec{k} - \vec{k}_0$, $\vec{k} = \vec{k}' + \vec{k}_0$ придем к выражению

$$\begin{aligned} G(\vec{x}, \vec{x}') &= \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \frac{\exp[i(\vec{k}', \vec{x}' - \vec{x})] \exp[i(\vec{k}_0, \vec{x}' - \vec{x})]}{-k'^2 - 2(\vec{k}', \vec{k}_0) - k_0^2 + 2k_0^2 + 2(\vec{k}_0, \vec{k}')} dk' = \\ &= \frac{\exp[i(\vec{k}_0, \vec{x}' - \vec{x})]}{(2\pi)^3} \iiint_{-\infty}^{\infty} \frac{\exp[i(\vec{k}', \vec{x}' - \vec{x})]}{k_0^2 - k'^2} dk'. \end{aligned}$$

Отметим, что последний интеграл является обратным преобразование Фурье для хорошо известной функции Грина для уравнения Гельмгольца:

$$\frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \frac{\exp[-i(\vec{k}', \vec{x}' - \vec{x})]}{k_0^2 - k'^2} dk' = -\frac{1}{(4\pi)} \frac{\exp[i|\vec{k}_0||\vec{x} - \vec{x}'|]}{|\vec{x} - \vec{x}'|}.$$

Используя это выражение получаем

$$\begin{aligned} G(\vec{x}, \vec{x}') &= \exp[-i(\vec{k}_0, \vec{x} - \vec{x}')] \left(\frac{-1}{4\pi} \frac{\exp[i|\vec{k}_0||\vec{x} - \vec{x}'|]}{|\vec{x} - \vec{x}'|} \right) = \\ &= \frac{-1}{4\pi} \frac{1}{|\vec{x} - \vec{x}'|} \exp[i|\vec{k}_0||\vec{x} - \vec{x}'|(1 - \cos \theta)], \end{aligned}$$

где θ – угол между векторами \vec{k}_0 и $\vec{x} - \vec{x}'$. Теперь можно записать первое приближение для рассеянного поля

$$\begin{aligned} u_1(\vec{x}, \omega) &= \iiint_{-\infty}^{\infty} G(\vec{x}, \vec{x}') \tilde{\beta}(\vec{x}') d\vec{x}' = \\ &= -\frac{1}{4\pi} \iiint_{-\infty}^{\infty} \frac{\tilde{\beta}(\vec{x}', \omega) \exp[i|\vec{k}_0||\vec{x} - \vec{x}'|(1 - \cos \theta)]}{|\vec{x} - \vec{x}'|} d\vec{x}'. \end{aligned}$$

Представим волновое поле в приближении Борна

$$\begin{aligned} u_B(\vec{x}, \omega) &= u_0(\vec{x}, \omega)(1 + \varepsilon u_1(\vec{x}, \omega)) = \\ &= a_0 \exp[i(\vec{k}_0, \vec{x})] + a_0 Q(\vec{x}, \omega, \vec{k}_0), \end{aligned}$$

где

$$\begin{aligned} Q(\vec{x}, \omega, \vec{k}_0) &= -\frac{\exp[i(\vec{k}_0, \vec{x})]}{4\pi} \iiint_{-\infty}^{\infty} \frac{\beta(\vec{x}, \omega)}{|\vec{x} - \vec{x}'|} \times \\ &\times \exp[i|\vec{k}_0||\vec{x} - \vec{x}'|(1 - \cos \theta)] d\vec{x}'. \end{aligned}$$

Приближение Борна применимо в случае слабых рассеивателей (слабоконтрастных возмущений среды), т.е. параметр $\beta(x, \omega)$ должен быть достаточно мал.

10.2 Модель сейсмической трассы в приближении Борна

Разрешающая способность дифракционной томографии может быть порядка длины волны или меньше. Приближение Борна может использоваться для восстановления пространственной скоростной функции или величины $\beta_0(x) = 1 - v_0^2/v^2(x)$ для слабо рассеивающих объектов.

Применимость приближения Борна определяется условием

$$\frac{\delta v}{v} \frac{d}{\lambda} \ll 1,$$

где $\delta v/v$ – относительная скоростная контрастность, d – размер области рассеяния, λ – длина волны зондирующего сигнала.

Используем приближение Борна для построения модели сейсмограмма

$$\tilde{u}(t, \vec{x}_r, \vec{n}) = u_{mB}(t, \vec{x}_r, \vec{n}, \beta_0(\vec{x})) + \xi(t, \vec{x}_r, \vec{n}).$$

Где u_{mB} – рассеянное поле в приближении Борна

$$u_{mB}(t, \vec{x}_r, \vec{n}, \beta_0(\vec{x})) = h(t) * F_{\omega}^{-1}[u_B(\vec{x}, \omega, \beta_0(\vec{x}))],$$

где $h(t)$ – аппаратная функция измерительного канала, F_{ω}^{-1} – оператор обратного преобразования Фурье,

$$u_B(\vec{x}, \omega) = a_0 \exp[i(\vec{k}_0, \vec{x})] + a_0 Q(\vec{x}_r, \omega, \vec{k}_0).$$

Основные определения, относящиеся к применению приближения Борна, представлены на рис. 10.3. Функцию $Q(\vec{x}_r, \omega, \vec{k}_0)$ в этом случае может быть записана

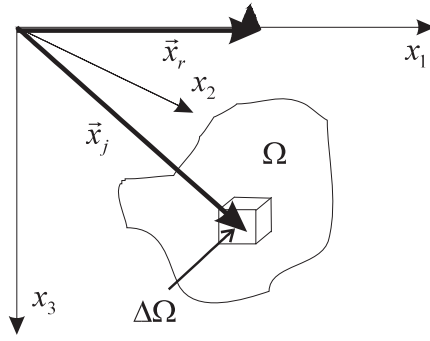


Рис. 10.3: Основные определения, относящиеся к применению приближения Борна.

в виде

$$\begin{aligned}
Q(\vec{x}_r, \omega, \vec{k}_0) &= -\frac{1}{4\pi} \iiint_{-\infty}^{\infty} \frac{\beta_0(\vec{x}) \omega^2}{v_0^2 |\vec{x} - \vec{x}'|} \exp[i|\vec{k}_0| |\vec{x}_r - \vec{x}'|] \times \\
&\times \exp[i(\vec{k}_0, \vec{x})] \exp[-i(\vec{k}_0, \vec{x}_r - \vec{x}')] d\vec{x}' = \\
&= \frac{-1}{4\pi v_0^2} \iiint_{-\infty}^{\infty} \frac{\beta_0(\vec{x}') \omega^2}{|\vec{x}_r - \vec{x}'|} \exp\{i[|\vec{k}_0| |\vec{x}_r - \vec{x}'| + (\vec{k}_0, \vec{x}')] \} d\vec{x}' = \\
&= \frac{1}{4\pi v_0^2} \iiint_{-\infty}^{\infty} \frac{(i\omega)^2}{|\vec{x}_r - \vec{x}'|} \beta_0(\vec{x}) \exp\left\{i\omega \left[\frac{|\vec{x}_r - \vec{x}'|}{v_0} + \frac{(\vec{n}, \vec{x}')}{v_0} \right] \right\} d\vec{x}'.
\end{aligned}$$

Представим неизвестную функцию $\beta_0(\vec{x})$ в виде разложения по базисным функциям

$$\beta_0(\vec{x}) = \left(1 - \frac{v_0^2}{v^2(\vec{x})}\right) = \sum_{j=1}^M \alpha_j \varphi_j(\vec{x}),$$

где

$$\varphi_j(\vec{x}) = \begin{cases} 1 & \text{для } \vec{x} \in \Delta\Omega_j, \\ 0 & \text{для } \vec{x} \notin \Delta\Omega_j. \end{cases}$$

Подставляя разложение для $\beta_0(\vec{x})$ в подынтегральное выражение, получаем

$$Q(\vec{x}_r, \omega, \vec{k}_0) = (i\omega)^2 \sum_{j=1}^M \alpha_j A_j \exp[i\omega\tau_j],$$

где

$$\tau_j = \left[\frac{|\vec{x}_r \vec{x}_j|}{v_0} + \frac{(\vec{n}, \vec{x}_j)}{v_0} \right], \quad A_j = \frac{\Delta\Omega_j}{4\pi v_0^2 |\vec{x}_r - \vec{x}_j|}.$$

Используя выражение для $Q(\vec{x}_r, \omega, k_0)$ можно переписать u_{mB} в виде

$$u_{mB} = a_0 h(t) * F_\omega^{-1} \left[\exp[i(\vec{k}_0, \vec{x}_r)] \right] + a_0 h(t) * F_\omega^{-1} \left[(i\omega)^2 \sum_{j=1}^M \exp[i\omega\tau_j] A_j \alpha_j \right] =$$

$$\begin{aligned}
&= a_0 h(t) * \delta \left(t - \frac{(\vec{n}, \vec{x}_r)}{v_0} \right) + a_0 h(t) * \sum_{j=1}^M \delta^{(2)}(t - \tau_j) \alpha_j A_j = \\
&= a_0 h \left(t - \frac{(\vec{n}, \vec{x}_r)}{v_0} \right) + a_0 \sum_{j=1}^M \alpha_j A_j h''(t - \tau_j).
\end{aligned}$$

Модель для $\tilde{u}(t, \vec{x}_r, \vec{n})$ можно переписать с учетом выражения полученного для $u_{mB}(t, \vec{x}_r, \vec{n}, \beta_0(\vec{x}))$:

$$u(t, \vec{x}_r, \vec{n}) = \sum_{j=1}^M \alpha_j C_j(t, \vec{x}_r, \vec{n}) + \varepsilon(t, \vec{x}_r, \vec{n}),$$

где

$$u(t, \vec{x}_r, \vec{n}) = \tilde{u}(t, \vec{x}_r, \vec{n}) - a_0 h(t - (\vec{n}, \vec{x}_r)/v_0), \quad C_j(t, \vec{x}_r, \vec{n}) = A_j h(t - \tau_j).$$

Так как сейсмограмма регистрируется в дискретных точках x_{ri} ($i = 1, 2, \dots, n$), запишем модель в матричной форме

$$\vec{u} = C\vec{\alpha} + \vec{\varepsilon},$$

где $\vec{u} = \|u(t, x_{ri}, \vec{n})\|_{i=1}^n$ – вектор измерений, при этом $C = \|C_j(t, x_{ri}, \vec{n})\|_{i=1, j=1}^{n \times M}$. Для оценивания вектора параметров $\vec{\alpha}$ мы можем использовать различные статистические критерии, основанные, например, методе статистической регуляризации или сингулярном анализе.

10.3 Дифракционная томография в случае векторного поля смещений

10.3.1 Модель измерений

Неизвестное поле параметров (параметры Ламе $\lambda(\vec{x})$, $\mu(\vec{x})$ и массовая плотность $\rho(\vec{x})$, или скорости p - и s -волн) являются элементами $\theta(\vec{x})$ функционального пространства $\Phi(R^3)$

$$\theta(\vec{x}) \in \Phi(\vec{x} \in R^3) \Rightarrow (\lambda(\vec{x}), \mu(\vec{x}), \rho(\vec{x})).$$

Пространство измерений – n -мерное Евклидово пространство (R^n) определяется отображением функционального пространства в пространство измерений

$$\Phi(\vec{x} \in R^3) \xrightarrow{\mathcal{P}} R^n$$

Где, экспериментальные данные $\{u(x_i, t_i), i = 1 \div n\} \in R^n$, $x \in R^3$ (3-компонентные сейсмограммы, содержащие шумовую компоненту) являются n -мерными функционалами искомого поля параметров $\theta(\vec{x})$

$$u_i = \mathcal{P}_i(\theta) + \varepsilon_i, \quad \varepsilon_i \in N(0, K_\varepsilon),$$

где $\mathcal{P}_i(\theta)$ оператор перехода из функционального пространства $\Phi(\vec{x} \in R^3)$ к пространству измерений R^n , ε_i – нормально распределенный шум с ковариационной матрицей K_ε математическим ожиданием 0. Томографический эксперимент состоит в регистрации зондирующего сигнала возбужденного источником f , и прошедшего через исследуемую среду. Пусть процесс распространения описывается линейным оператором L_θ

$$L_\theta \varphi = f,$$

$$L_\theta : L_\theta(\alpha\varphi + \beta\psi) = \alpha L_\theta \varphi + \beta L_\theta \psi.$$

Преобразование зондирующего сигнала приемным каналом описывается интегральным линейным оператором свертки

$$H : H_i \varphi \leftrightarrow \int_T \int_\Omega \int_V d\vec{x} d\tau d\Omega h_i(\vec{n}, \vec{n}', t - \tau) \varphi(\vec{x}, \vec{n}, \tau) \delta(\vec{x} - \vec{x}_i),$$

где Ω – телесный угол, \vec{n}' – направляющий вектор главного лепестка диаграммы направленности приемника – см. рис. 10.4)).

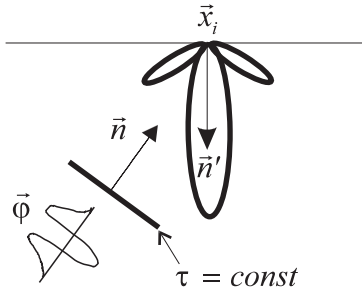


Рис. 10.4: Падающее поле и диаграмма направленности приемника.

В обобщенном виде модель экспериментальных данных может быть записана следующим образом

$$u_i = H_i L_\theta^{-1} \vec{f} + \varepsilon_i.$$

В этом случае оператор перехода:

$$\mathcal{P}_i(\theta) = H_i L_\theta^{-1} \vec{f}.$$

В действительности невозможно получить точное решение задачи о нахождении параметров среды по экспериментальным данным. Будем использовать для этого приближенный метод. Предположим, что для некоторых значений θ_0 можно построить решение

$$\varphi_0 = L_0^{-1} f.$$

Предположим так же, что искомое значение θ не слишком сильно отличается от θ_0 , i.e.

$$\theta = \theta_0 + \delta\theta,$$

где $\delta\theta \ll \theta_0$.

Формальное решение для φ определяется выражением

$$\varphi = \varphi_0 + L_0^{-1} \delta L_\theta \varphi,$$

где $\delta L_\theta = L_0 - L_\theta$ – возмущающий оператор. Можно переписать модель в виде

$$u_i = H_i[\varphi_0 + L_0^{-1} \delta L_\theta \varphi] + \varepsilon_i.$$

Модель нелинейна относительно $\delta\theta$. Если выполнено условие

$$\frac{\|H_i L_0^{-1} \delta L_\theta (\varphi - \varphi_0)\|^2}{E(\varepsilon_i^2)} \ll 1$$

(модельная ошибка много меньше ошибки измерения), то величину φ можно заменить на φ_0

$$u_i = H_i[\varphi_0 + L_0^{-1} \delta L_\theta \varphi_0] + \tilde{\varepsilon}_i,$$

где $\tilde{\varepsilon}_i$ включает как случайную ошибку ε_i , так и ошибку в задании модели. Перепишем модель введя скалярное произведение

$$\tilde{u}_i = \langle h_i | L_0^{-1} \delta L_\theta \varphi_0 \rangle_{V,T,\Omega} + \tilde{\varepsilon}_i$$

где

$$\langle \xi | \eta \rangle_{T,V,\Omega} = \int_{\Omega} \int_V \xi(\vec{x}, t, \Omega) * \eta(\vec{x}, t, \Omega) d\vec{x} d\Omega$$

– скалярное произведение, $*$ – обозначение свертки по времени,

$$\tilde{u}_i = u_i - \langle h_i | \varphi_0 \rangle_{T,V,\Omega}$$

– разность между зарегистрированными данными u_i и полученными в результате решения прямой задачи с параметрами $\theta_0(\vec{x})$.

10.3.2 Томографический функционал

Выделим в возмущающем операторе δL монотонную функцию $\nu(\delta\theta)$ относительно которой возмущающий оператор линеен

$$\begin{aligned} \tilde{u}_i &= \langle L_0^{-1*} h_i | \delta L_\theta \varphi_0 \rangle_{VT\Omega} + \tilde{\varepsilon}_i = \\ &= \langle \langle G^* h_i | L'_\nu \varphi_0 \rangle_{T\Omega} | \nu(\delta\theta) \rangle_V + \tilde{\varepsilon}_i, \end{aligned}$$

где $G = L_0^{-1}$ – оператор Грина, G^* – сопряженный оператор Грина, L'_ν – производная Фреше:

$$L'_\nu : \frac{\delta u}{\delta \nu} = \langle G^* h | L'_\nu \varphi_0 \rangle_{T\Omega}.$$

Интегральное ядро функционала измерений относительно $\nu(\delta\theta)$ носит название томографического функционала

$$p_i = \langle \varphi_{out} | S | \varphi_{in} \rangle_{T\Omega},$$

где

$\varphi_{in} = \varphi_0 = Gf$ – падающее поле в известной опорной среде θ_0 ,

$\varphi_{out} = G^* h_i$ – отраженное поле (когда точкой излучения является точка приема),

$S = L'_\nu$ – оператор взаимодействия полей φ_{in} и φ_{out} .

Воспользовавшись определением томографического функционала, запишем модель в виде

$$\tilde{u} = P\nu + \tilde{\varepsilon},$$

где $\tilde{u} = \|\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n\|^T$ – вектор наблюдаемых величин, $\tilde{\varepsilon} = \|\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \dots, \tilde{\varepsilon}_n\|^T$ – случайный вектор ошибок,

$$\nu = \left\| \begin{array}{c} |\nu_1\rangle \\ \dots \\ |\nu_M\rangle \end{array} \right\|$$

– вектор функционалов от неизвестных параметров,

$$P = \left\| \begin{array}{cccc} \langle p_{11}| & \dots & \langle p_{1\mu}| & \dots & \langle p_{1M}| \\ \dots & \dots & \dots & \dots & \dots \\ \langle p_{n1}| & \dots & \langle p_{n\mu}| & \dots & \langle p_{nM}| \end{array} \right\|$$

– матрица томографических функционалов.

Томографический функционал p_i определяет влияние всех элементов пространственной области на формирование величин u_i наблюдаемых в эксперименте

$$u_i = \langle p_i | \nu(\delta\theta) \rangle_V.$$

В традиционной лучевой томографии томографический функционал

$$p_i = \delta(L_i(\vec{x}))$$

является сингулярным и локализуется вдоль луча, соединяющего источник и приемник (см. рис. 10.5).

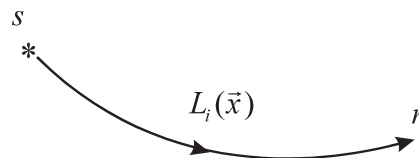


Рис. 10.5: В лучевой томографии томографический функционал локализуется вдоль луча, соединяющего источник и приемник.

В дифракционной томографии каждый элемент объема обладает своим собственным весом, который определяется характером взаимодействия полей φ_{in} и φ_{out} . Основное содержание томографического эксперимента связано с перекрытием носителей томографических функционалов.

10.3.3 Томографический функционал для скалярного волнового уравнения

Оператор, отвечающий за распространение в опорной среде:

$$L_0 = -\Delta + \frac{1}{v_0^2(x)} \frac{\partial^2}{\partial t^2}.$$

Оператор, отвечающий за распространение в искомой (неизвестной) среде:

$$L_\theta = -\Delta + \frac{1}{v^2(x)} \frac{\partial^2}{\partial t^2},$$

где $\theta_0 = v_0$, $\theta = v(x)$, $x \in R^3$,

$$\nu = \frac{1}{v_0^2} \left(1 - \frac{v_0^2}{v^2(x)} \right)$$

– неизвестная (искомая) функция,

$$S = -\frac{\partial^2}{\partial t^2}$$

– оператор взаимодействия,

$$p = \langle \varphi_{out} | \frac{\partial^2}{\partial t^2} | \varphi_{in} \rangle_T$$

– томографический функционал.

Носитель томографического функционала в однородной опорной среде локализуется в параболическом слое, когда падающее поле является плоской волной, и в эллиптическом слое, когда падающее поле является сферической волной (при условии точечности приемника) (см. рис. 10.6).

10.3.4 Томографический функционал для уравнения Ламе

Оператор, отвечающий за распространение в опорной среде:

$$\begin{aligned} L_0 \vec{\varphi} = & \rho_0 \frac{\partial^2}{\partial t^2} \vec{\varphi} - [(\lambda_0 + \mu_0) \vec{\nabla} \vec{\nabla} \cdot \vec{\varphi} + \mu_0 \Delta \vec{\varphi} + \vec{\nabla} \lambda_0 \vec{\nabla} \cdot \vec{\varphi} + \\ & + \vec{\nabla} \mu_0 \times \vec{\nabla} \times \vec{\varphi} + 2(\vec{\nabla} \mu_0 \cdot \vec{\nabla}) \vec{\varphi}]. \end{aligned}$$

Оператор, отвечающий за распространение в искомой среде:

$$\begin{aligned} L \vec{\varphi} = & \rho \frac{\partial^2}{\partial t^2} \vec{\varphi} - [(\lambda + \mu) \vec{\nabla} \vec{\nabla} \cdot \vec{\varphi} + \mu \Delta \vec{\varphi} + \vec{\nabla} \lambda \vec{\nabla} \cdot \vec{\varphi} + \\ & + \vec{\nabla} \mu \times \vec{\nabla} \times \vec{\varphi} + 2(\vec{\nabla} \mu \cdot \vec{\nabla}) \vec{\varphi}]. \end{aligned}$$

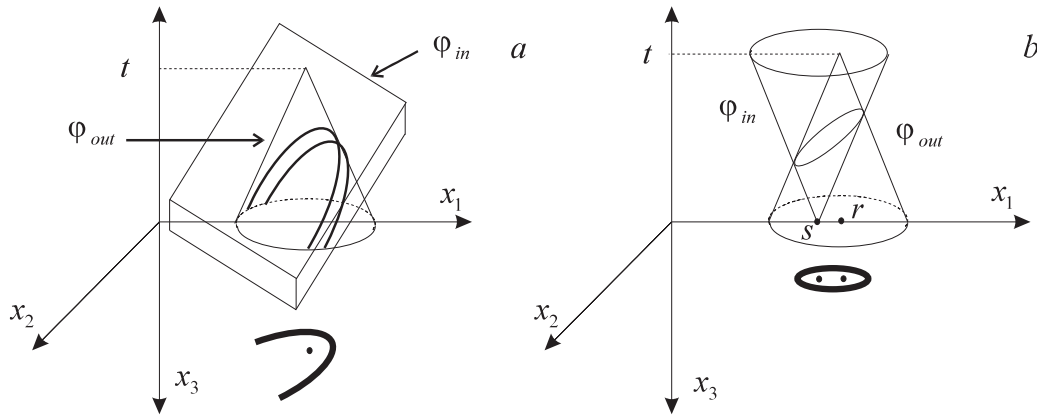


Рис. 10.6: Графическое представление взаимодействия между φ_{in} и φ_{out} . а – плоская падающая волна; б – сферическая падающая волна.

Величины θ_0 и θ

$$\theta_0 = \left\| \begin{array}{l} |\lambda_0(x)\rangle \\ |\mu_0(x)\rangle \\ |\rho_0(x)\rangle \end{array} \right\| \quad \theta = \left\| \begin{array}{l} |\lambda(x)\rangle \\ |\mu(x)\rangle \\ |\rho(x)\rangle \end{array} \right\|$$

векторы параметров в опорной среде и искомой среде соответственно, где $\lambda(x)$ и $\mu(x)$ – параметры Ламе, $\rho(x)$ – массовая плотность. Вектор томографических функционалов и оператор взаимодействия с блоковой структурой соответственно записываются в виде

$$p = \|\langle p_\lambda | \langle p_\mu | \langle p_\rho | \| \quad \text{и} \quad S = \|S_\lambda | S_\mu | S_\rho \|.$$

Возмущающий оператор равен

$$\delta L_\theta = L_0 - L = \delta L_\lambda + \delta L_\mu + \delta L_\rho.$$

Представим неизвестные параметры как возмущения опорной среды

$$\lambda(x) = \lambda_0(x) + \delta\lambda(x), \quad \mu(x) = \mu_0(x) + \delta\mu(x), \quad \rho(x) = \rho_0(x) + \delta\rho(x).$$

Явная запись возмущающего оператора δL_λ :

$$\delta L_\lambda : \delta L_\lambda \vec{\varphi} = \delta\lambda \vec{\nabla} \vec{\nabla} \cdot \vec{\varphi} + \vec{\nabla} \delta\lambda \vec{\nabla} \cdot \vec{\varphi} = \vec{\nabla} (\delta\lambda \vec{\nabla} \cdot \vec{\varphi}),$$

при этом используется равенство $\vec{\nabla}(\phi\psi) = \psi \vec{\nabla}\phi + \phi \vec{\nabla}\psi$ используется. Явная запись возмущающего оператора для δL_μ :

$$\begin{aligned} \delta L_\mu : \quad \delta L_\mu \vec{\varphi} = & \delta\mu \vec{\nabla} \vec{\nabla} \cdot \vec{\varphi} + \delta\mu \Delta \vec{\varphi} + \vec{\nabla} \delta\mu \times \vec{\nabla} \times \vec{\varphi} + \\ & + 2(\vec{\nabla} \delta\mu \cdot \vec{\nabla}) \vec{\varphi} = 2\vec{\nabla} \cdot (\delta\mu \vec{\nabla} \vec{\varphi}) + \vec{\nabla} \times (\delta\mu \vec{\nabla} \times \vec{\varphi}). \end{aligned}$$

Последнее отношение получено с использованием равенства

$$2\vec{\nabla} \cdot (\delta\mu \vec{\nabla} \vec{\varphi}) = 2(\vec{\nabla} \delta\mu \cdot \vec{\nabla}) \vec{\varphi} + 2\delta\mu \Delta \vec{\varphi},$$

которое является следствием тождества

$$\vec{\nabla} \cdot (\varphi \vec{F}) = \vec{\nabla} \varphi \cdot \vec{F} + \varphi \vec{\nabla} \cdot \vec{F}$$

и равенства

$$\begin{aligned} \vec{\nabla} \times (\delta\mu \underbrace{\vec{\nabla} \times \vec{\varphi}}) &= \vec{\nabla} \delta\mu \times \vec{\nabla} \times \vec{\varphi} + \delta\mu \vec{\nabla} \times \vec{\nabla} \times \vec{\varphi} = \\ &= \vec{\nabla} \delta\mu \times \vec{\nabla} \times \vec{\varphi} + \delta\mu \vec{\nabla} \vec{\nabla} \cdot \vec{\varphi} - \delta\mu \Delta \vec{\varphi} \end{aligned}$$

которое следует из тождеств

$$\begin{aligned} \vec{\nabla} \times (\phi \vec{F}) &= \vec{\nabla} \phi \times \vec{F} + \phi \vec{\nabla} \times \vec{F}, \\ \vec{\nabla} \times \vec{\nabla} \times \vec{F} &= \vec{\nabla} \vec{\nabla} \cdot \vec{F} - \Delta \vec{F}. \end{aligned}$$

Явная запись возмущающего оператора δL_ρ :

$$\delta L_\rho : \delta L_\rho \vec{\varphi} = -\delta\rho \partial^2 \vec{\varphi} / \partial t^2.$$

С учетом представления для томографических функционалов, можно записать модель измерений в виде

$$\vec{u}_i = u_i - u_{i0} = \langle p_{i\lambda} | \delta\lambda \rangle_V + \langle p_{i\mu} | \delta\mu \rangle_V + \langle p_{i\rho} | \delta\rho \rangle_V + \varepsilon_i = \langle p_{i\theta} | \delta\theta \rangle_V + \varepsilon_i,$$

где используется следующее обозначение

$$\langle p_{i\theta} | \delta\theta \rangle_V = \int_V \vec{\varphi}_{out} \odot \delta L_\theta \vec{\varphi}_{in} d\vec{x} = \langle \langle \vec{\varphi}_{out} | S_\theta | \vec{\varphi}_{in} \rangle_T | \delta\theta \rangle_V.$$

Знак \odot принят для совместного обозначения операции свертки $*$ и скалярного произведения (\cdot, \cdot)

$$\vec{\xi} \odot \vec{\eta} \equiv (\vec{\xi}(t) * \vec{\eta}(t)).$$

Используя запись возмущающего оператора, полученную ранее, запишем оператор взаимодействия в явном виде

$$\begin{aligned} S_\lambda : \int_V \vec{\varphi}_{out} \odot \delta L_\lambda \vec{\varphi}_{in} d\vec{x} &= \int_V \vec{\varphi}_{out} \odot \vec{\nabla} (\delta\lambda \vec{\nabla} \cdot \vec{\varphi}_{in}) d\vec{x} = \\ &= \int_V \vec{\nabla} \cdot (\vec{\varphi}_{out} * \delta\lambda \vec{\nabla} \cdot \vec{\varphi}_{in}) d\vec{x} - \int_V \vec{\nabla} \cdot \vec{\varphi}_{out} * \delta\lambda \vec{\nabla} \cdot \vec{\varphi}_{in} d\vec{x}, \end{aligned}$$

используя тождество

$$\vec{\nabla} (\phi \vec{\xi}) = \vec{\xi} \cdot \vec{\nabla} \phi + \phi \vec{\nabla} \cdot \vec{\xi}.$$

Обращаясь к теореме Остроградского-Гаусса получаем

$$\int_V \vec{\nabla} \cdot (\vec{\varphi}_{out} * \delta\lambda \vec{\nabla} \cdot \vec{\varphi}_{in}) d\vec{x} = \int_{\partial V} (d\vec{S} \cdot \vec{\varphi}_{out} * \delta\lambda \vec{\nabla} \cdot \vec{\varphi}_{in}) = 0.$$

Выбираем объем интегрирования достаточно большим, таким что $\delta\lambda|_{\partial V} \equiv 0$ и поверхностный интеграл обращается в ноль. Окончательно получаем явную запись томографического функционала p_λ

$$S_\lambda : p_\lambda = \langle \vec{\varphi}_{out} | S_\lambda | \vec{\varphi}_{in} \rangle_T = -\vec{\nabla} \cdot \vec{\varphi}_{out} * \vec{\nabla} \cdot \vec{\varphi}_{in}.$$

Оператор взаимодействия S_μ можно записать в общем виде

$$S_\mu : 2 \int_V \vec{\varphi}_{out} \odot \vec{\nabla} \cdot (\delta\mu \vec{\nabla} \vec{\varphi}_{in}) d\vec{x} + \int_V \vec{\varphi}_{out} \odot \vec{\nabla} \times (\delta\mu \vec{\nabla} \vec{\varphi}_{in}) d\vec{x},$$

используя тождества

$$\vec{\nabla} \vec{\varphi} \oplus \vec{\nabla} \vec{\psi} = \sum_{i=1}^3 \sum_{j=1}^3 \nabla_i \varphi_j * \nabla_i \psi_j$$

$$\vec{\varphi} \otimes \vec{\nabla} \vec{\psi} = \sum_{i=1}^3 \varphi_i * \nabla_i \psi_j, \quad \vec{\nabla} \cdot (\vec{\xi} \cdot \mathbf{T}) = \vec{\nabla} \vec{\xi} \cdot \mathbf{T} + \vec{\xi} \cdot \vec{\nabla} \cdot \mathbf{T},$$

получаем для первого интеграла

$$\begin{aligned} & 2 \int_V \vec{\varphi}_{out} \odot \vec{\nabla} \cdot (\delta\mu \vec{\nabla} \vec{\varphi}_{in}) d\vec{x} = \\ & = 2 \int_V \vec{\nabla} \cdot (\vec{\varphi}_{out} \otimes \delta\mu \vec{\nabla} \vec{\varphi}_{in}) d\vec{x} - 2 \int_V \vec{\nabla} \vec{\varphi}_{out} \oplus \delta\mu \vec{\nabla} \vec{\varphi}_{in} d\vec{x} = \\ & = -2 \int_V \vec{\nabla} \vec{\varphi}_{out} \oplus \vec{\nabla} \vec{\varphi}_{in} \delta\mu d\vec{x}. \end{aligned}$$

Поскольку по теореме Остроградского-Гаусса первый интеграл обращается в нуль вследствие выполнения равенства $\delta\mu|_{\partial V} = 0$

$$\int_V \vec{\nabla} \cdot (\vec{\varphi}_{out} \otimes \delta\mu \vec{\nabla} \vec{\varphi}_{in}) d\vec{x} = \int_{\partial V} d\vec{S} \cdot (\vec{\varphi}_{out} \otimes \delta\mu \vec{\nabla} \vec{\varphi}_{in}) = 0.$$

Второй интеграл можно записать в виде

$$\begin{aligned} & \int_V \vec{\varphi}_{out} \odot \vec{\nabla} \times (\delta\mu \vec{\nabla} \times \vec{\varphi}_{in}) d\vec{x} = \\ & = - \int_V \vec{\nabla} \cdot (\vec{\varphi}_{out} \otimes (\delta\mu \vec{\nabla} \times \vec{\varphi}_{in})) d\vec{x} + \\ & + \int_V (\delta\mu \vec{\nabla} \times \vec{\varphi}_{in}) \odot (\vec{\nabla} \times \vec{\varphi}_{out}) d\vec{x} = \\ & = \int_V \delta\mu (\vec{\nabla} \times \vec{\varphi}_{out}) \odot (\vec{\nabla} \times \vec{\varphi}_{in}) d\vec{x}, \end{aligned}$$

где используется тождество

$$\vec{\nabla} \cdot (\vec{\xi} \times \vec{\eta}) = \vec{\eta} \cdot \vec{\nabla} \times \vec{\xi} - \vec{\xi} \cdot \vec{\nabla} \times \vec{\eta}, \quad \vec{\eta} = \vec{\varphi}_{out}, \quad \vec{\xi} = (\delta\mu \vec{\nabla} \times \vec{\varphi}_{in})$$

и обозначение

$$\otimes : \vec{\xi} \otimes \vec{\eta} = \int \vec{\xi}(\tau) \times \vec{\eta}(t - \tau) d\tau.$$

По теореме Остроградского-Гаусса первый интеграл обращается в нуль

$$\begin{aligned} \int_V \vec{\nabla} \cdot (\vec{\varphi}_{out} \otimes (\delta\mu \vec{\nabla} \times \vec{\varphi}_{in})) d\vec{x} &= \\ &= \int_{\partial V} d\vec{S} \cdot (\vec{\varphi}_{out} \otimes (\delta\mu \vec{\nabla} \times \vec{\varphi}_{in})) = 0 \end{aligned}$$

вследствие выполнения условия $\delta\mu|_{\partial V} = 0$. С учетом полученного представления для оператора взаимодействия S_μ , запишем возмущающий оператор δL_μ в виде

$$\begin{aligned} \delta L_\mu : \int_V \vec{\varphi}_{out} \odot \delta L_\mu \vec{\varphi}_{in} d\vec{x} &= \\ &= \int_V [(\vec{\nabla} \times \vec{\varphi}_{out}) \odot (\vec{\nabla} \times \vec{\varphi}_{in}) - 2\vec{\nabla} \vec{\varphi}_{out} \oplus \vec{\nabla} \vec{\varphi}_{in}] \delta\mu d\vec{x} = \\ &= \int_V p_\mu(\vec{x}) \delta\mu(\vec{x}) d\vec{x} = \langle \langle \vec{\varphi}_{out} | S_\mu | \vec{\varphi}_{in} \rangle_T | \delta\mu \rangle_V. \end{aligned}$$

Запишем в окончательном виде томографический функционал p_μ и оператор взаимодействия S_μ

$$\boxed{S_\mu : p_\mu = \langle \vec{\varphi}_{out} | S_\mu | \vec{\varphi}_{in} \rangle_T = (\vec{\nabla} \times \vec{\varphi}_{out}) \odot (\vec{\nabla} \times \vec{\varphi}_{in}) - 2\vec{\nabla} \vec{\varphi}_{out} \oplus \vec{\nabla} \vec{\varphi}_{in}.}$$

Действие оператора S_ρ определяется выражением

$$S_\rho : \int_V \vec{\varphi}_{out} \odot \delta L_\rho \vec{\varphi}_{in} d\vec{x} = - \int_V \vec{\varphi}_{out} \odot \delta\rho \frac{\partial^2}{\partial t^2} \vec{\varphi}_{in} d\vec{x}.$$

После интегрирования по частям получаем

$$\begin{aligned} \vec{\varphi}_{out} \odot \frac{\partial^2}{\partial t^2} \vec{\varphi}_{in} &= \int_{-\infty}^{\infty} \vec{\varphi}_{out}(t_i - t) \cdot \frac{\partial^2}{\partial t^2} \vec{\varphi}_{in}(t) dt = \\ &= \vec{\varphi}_{out} \cdot \frac{\partial}{\partial t} \vec{\varphi}_{in} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\partial}{\partial t} \vec{\varphi}_{out}(t_i - t) \cdot \frac{\partial}{\partial t} \vec{\varphi}_{in}(t) dt. \end{aligned}$$

Принимая во внимание условия

$$\varphi_{out} \Big|_{\infty} = 0, \quad \frac{\partial}{\partial t} \varphi_{in} \Big|_{-\infty} = 0,$$

окончательно томографический функционал p_ρ и оператор взаимодействия S_ρ можно записать в виде

$$S_\rho : p_\rho = \langle \vec{\varphi}_{out} | S_\rho | \vec{\varphi}_{in} \rangle_T = \frac{\partial}{\partial t} \vec{\varphi}_{out} \odot \frac{\partial}{\partial t} \vec{\varphi}_{in} .$$

Чтобы проиллюстрировать конфигурацию носителя томографического функционала, рассмотрим плоскую падающую волну. Характеристические конусы, соответствующие полю точечного источника $\vec{\varphi}_{out}$, генерируются приемником в момент t_i и распространяются в обратном времени. Внешняя коническая поверхность соответствует распространению со скоростью волны сжатия (p -волна) внутренней волны со скоростью волны сдвига (s -волна). Волновое поле $\vec{\varphi}_{in}$ распространяется в положительном направлении оси по оси x . Носитель томографического функционала ограничен в пространстве параболоидом вращения, образованным пересечением внешнего конуса $\vec{\varphi}_{out}$ и поля $\vec{\varphi}_{in}$ и эллипсоидом вращения, образованным пересечением внутреннего конуса $\vec{\varphi}_{out}$ и поля $\vec{\varphi}_{in}$ (см. рис. 10.7). Если падающее поле является s -волной, то носитель томографического функци-

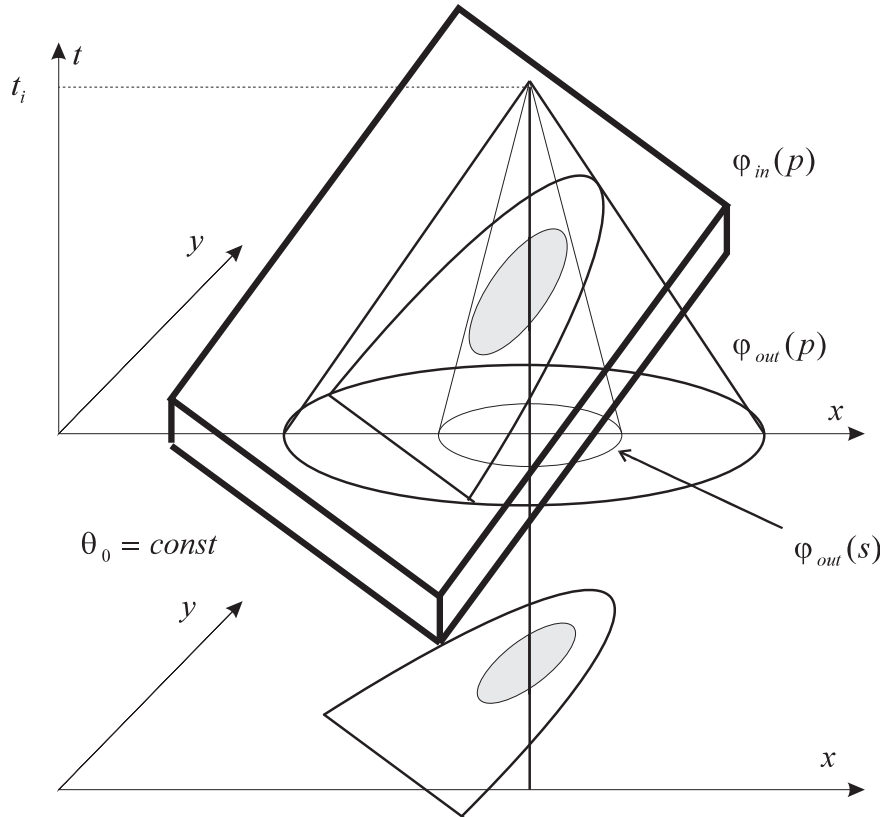


Рис. 10.7: Графическое представление взаимодействия между $\vec{\varphi}_{in}$ и $\vec{\varphi}_{out}$. $\vec{\varphi}_{in}$ — p -волна.

онала ограничен двумя параболоидами вращения (см. рис. 10.8).

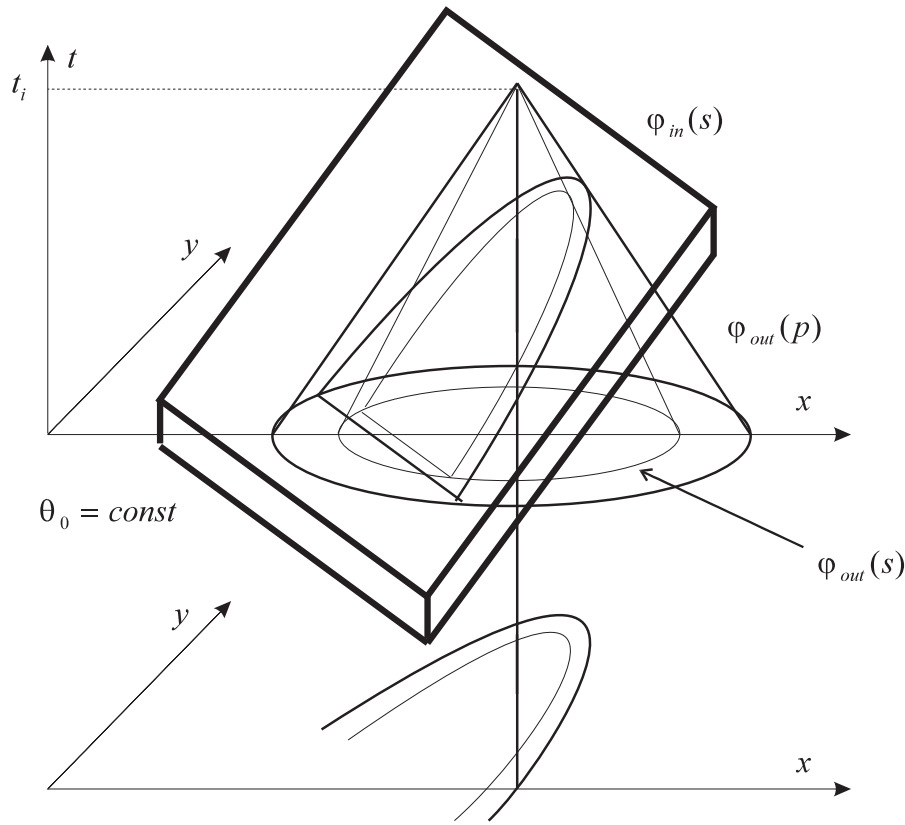


Рис. 10.8: Графическое представление взаимодействия $\vec{\varphi}_{in}$ и $\vec{\varphi}_{out}$. $\vec{\varphi}_{in}$ – s -волна.

10.4 Алгоритм реконструкции

Чтобы построить алгоритм реконструкции, запишем модель измерений используя понятие томографического функционала

$$u = P\nu + \varepsilon, \quad u_i = \sum_{\mu=1}^M \langle p_{\mu i} | \nu_{\mu}(\delta\theta) \rangle_V + \varepsilon_i.$$

Представим произвольный линейный функционал $l(\nu)$ как линейную комбинацию измеренных величин

$$l(\nu) = \sum_{i=1}^n \alpha_i u_i = (\vec{\alpha} \cdot \vec{u}),$$

где $l(\nu) \equiv \langle l | \nu \rangle_V$ – линейный функционал, например, средняя величина в элементарном объеме (сфера или параллелепипед – см. рис. 10.9):

$$l^X(\nu) = \left(\frac{4}{3} \pi r^3 \right)^{-1} H(r - |x - X|),$$

$$l^X(\nu) = \frac{1}{abc} H(a/2 - |x_1 - X_1|) H(b/2 - |x_2 - X_2|) H(c/2 - |x_3 - X_3|),$$

где

$$H(x) = \begin{cases} 1 & \text{для } x > 0, \\ 0 & \text{для } x < 0. \end{cases}$$

Графическое представление функции $H(x)$ – см. рис. 10.10. Независимо от кон-

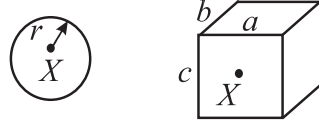


Рис. 10.9: Примеры элементарных объемов.

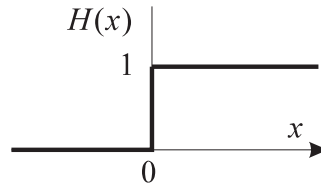


Рис. 10.10: Графическое представление функции $H(x)$.

кретного выбора процедуры, ошибку восстановления можно записать как

$$\begin{aligned} \eta_\mu^X &= l(\nu_\mu) - \sum_{\mu'=1}^M (\vec{\alpha} \cdot \langle \vec{p}_{\mu'} | \nu_{\mu'} \rangle_V) - (\vec{\alpha} \cdot \vec{\varepsilon}) = \\ &= \langle l - (\vec{\alpha} \cdot \vec{p}_\mu) | \nu_\mu \rangle_V - \sum_{\mu' \neq \mu} \langle (\vec{\alpha} \cdot \vec{p}_{\mu'}) | \nu_{\mu'} \rangle_V - (\vec{\alpha} \cdot \vec{\varepsilon}). \end{aligned}$$

В качестве решения возьмем функцию

$$\hat{l}(\nu) = (\vec{\alpha} \cdot \vec{u}),$$

которая минимизирует квадрат ошибки восстановления

$$E(\eta_\mu^X)^2 = (\vec{\Lambda} - Q\vec{\alpha})^* K (\vec{\Lambda} - Q\vec{\alpha}),$$

где $Q = \|P|I\|$ и I – единичный оператор с размерностью равной числу измерений

$$\begin{aligned} \alpha^T Q^* &= \|\alpha_1 \dots \alpha_n\| \times \\ &\times \left\| \begin{array}{cccccc} \langle p_{11}(x) | & \dots & \langle p_{1\mu}(x) | & \dots & \langle p_{1M}(x) | & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \langle p_{i1}(x) | & \dots & \langle p_{i\mu}(x) | & \dots & \langle p_{iM}(x) | & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \langle p_{n1}(x) | & \dots & \langle p_{n\mu}(x) | & \dots & \langle p_{nM}(x) | & 0 & \dots & 1 \end{array} \right\|, \\ \Lambda^* &= \left\| \begin{array}{cccc|ccc} \langle 0 & \dots & \langle l(x) & \dots & \langle 0 & 0 & \dots & 0 \\ 1 & & \mu & & M & M+1 & & M+n \end{array} \right\|. \end{aligned}$$

Матрица K – корреляционный оператор

$$K = \begin{vmatrix} K_{\nu_1\nu_1} & \cdots & K_{\nu_1\nu_M} & K_{\nu_1\varepsilon} \\ \cdots & \cdots & \cdots & \cdots \\ K_{\nu_M\nu_1} & \cdots & K_{\nu_M\nu_M} & K_{\nu_M\varepsilon} \\ K_{\varepsilon\nu_1} & \cdots & K_{\varepsilon\nu_M} & K_{\varepsilon\varepsilon} \end{vmatrix},$$

$$K : \langle \varphi | K_{\nu_\mu\nu_{\mu'}} | \psi \rangle = E \langle \varphi | \nu_\mu \rangle \langle \nu_{\mu'} | \psi \rangle,$$

$$K_{\varepsilon\varepsilon} = E(\vec{\varepsilon} \cdot \vec{\varepsilon}^T), \quad K_{\varepsilon\nu_\mu} : E(\varepsilon \langle \nu_\mu | \psi \rangle) = K_{\varepsilon\nu_\mu} | \psi \rangle.$$

Оптимальная оценка $\vec{\alpha}$ может быть представлена в виде

$$\hat{\vec{\alpha}} = (QKQ^*)^{-1}QK\vec{\Lambda}.$$

В частном случае, когда статистические связи между различными полями ν_μ и $\nu_{\mu'}$ отсутствуют, также как и между ν_μ и ε , оптимальная оценка может быть представлена в виде

$$\hat{\alpha}_\mu = (P_\mu K_{\nu_\mu\nu_\mu} P_\mu^* + \sum_{\mu' \neq \mu} P_{\mu'} K_{\nu_{\mu'}\nu_{\mu'}} P_{\mu'}^* + K_{\varepsilon\varepsilon})^{-1} \langle p_\mu K_{\nu_\mu\nu_\mu} | l \rangle,$$

где

$$\begin{aligned} P_\mu K_{\nu_\mu\nu_\mu} P_\mu^* &= \begin{vmatrix} \langle p_{\mu 1} | \\ \cdots \\ \langle p_{\mu n} | \end{vmatrix} \begin{vmatrix} K_{\nu_\mu\nu_\mu} \\ \cdots \\ \cdots \end{vmatrix} \begin{vmatrix} | p_{\mu 1} \rangle \\ \cdots \\ | p_{\mu n} \rangle \end{vmatrix} = \\ &= \begin{vmatrix} \langle p_{\mu 1}(x) | K_{\nu_\mu\nu_\mu}(x, x') | p_{\mu 1}(x') \rangle & \cdots & \langle p_{\mu 1}(x) | K_{\nu_\mu\nu_\mu}(x, x') | p_{\mu n}(x') \rangle \\ \cdots & \cdots & \cdots \\ \langle p_{\mu n}(x) | K_{\nu_\mu\nu_\mu}(x, x') | p_{\mu 1}(x') \rangle & \cdots & \langle p_{\mu n}(x) | K_{\nu_\mu\nu_\mu}(x, x') | p_{\mu n}(x') \rangle \end{vmatrix} \end{aligned}$$

– ковариационная матрица полезного (относительного восстанавливаемого μ -о поля параметров) сигнала,

$$\left(\sum_{\mu' \neq \mu} P_{\mu'} K_{\nu_{\mu'}\nu_{\mu'}} P_{\mu'}^* + K_{\varepsilon\varepsilon} \right)$$

– ковариационная матрица эффективного шума.

Запишем дисперсию ошибки оценки

$$\begin{aligned} E(\eta_\mu)^2 &= \langle l | K_{\nu_\mu\nu_\mu} - K_{\nu_\mu\nu_\mu} P_\mu^* (P_\mu K_{\nu_\mu\nu_\mu} P_\mu^* + \\ &+ \sum_{\mu' \neq \mu} P_{\mu'} K_{\nu_{\mu'}\nu_{\mu'}} P_{\mu'}^* + K_{\varepsilon\varepsilon})^{-1} P_\mu K_{\nu_\mu\nu_\mu} | l \rangle \end{aligned}$$

или, вводя оператор Фишера,

$$E(\eta_\mu)^2 = \langle l | F_\mu^{-1} | l \rangle,$$

где

$$F_\mu = [K_{\nu_\mu\nu_\mu} - K_{\nu_\mu\nu_\mu} P_\mu^* (P_\mu K_{\nu_\mu\nu_\mu} P_\mu^* + \sum_{\mu' \neq \mu} P_{\mu'} K_{\nu_{\mu'}\nu_{\mu'}} P_{\mu'}^* + K_{\varepsilon\varepsilon})^{-1} P_\mu K_{\nu_\mu\nu_\mu}]^{-1}.$$

Из последнего выражения следует, что качество оценки μ -о поля тем лучше, чем выше чувствительность обрабатываемых данных к вариациям μ -о поля и чем меньше влияние вариаций остальных полей на экспериментальные данные.

10.5 Анализ ошибки восстановления и сравнение алгоритма реконструктивной томографии с методом Бейкуса-Гильберта

Искажение измеряемых величин случайным шумом фактически приводит к тому, что процедура оценивания полей $\nu_\mu(\delta\theta)$ является процедурой статистического оценивания.

Запишем ошибку оценивания одного поля $\nu(\delta\theta)$

$$\eta_l = \underbrace{\langle l - (\vec{\alpha} \cdot \vec{p}) | \nu \rangle}_b - \underbrace{(\vec{\alpha} \cdot \vec{\varepsilon})}_n.$$

Ошибка η_l состоит из двух частей: смещения (b) и шума (n).

Рассмотрим критерий оптимальности в смысле критерия Бейкуса-Гильберта (BG), в котором декларируется отсутствие априорной информации относительно поля параметров $\nu(\delta\theta)$.

Первая часть критерия W_{BG} – “дельтообразность” L т.е. оценивается сингулярный функционал

$$l_{BG}^x = \delta(x - X)$$

и, если обозначить

$$\hat{l} \equiv (\vec{\alpha}_x \cdot \vec{p}) = (\vec{\alpha}(X) \cdot \vec{p}(x)),$$

то

$$L = \langle l_{BG} - \hat{l} | K_{BG}^X | l_{BG} - \hat{l} \rangle,$$

где интегральный оператор K_{BG}^X имеет сингулярное ядро вида

$$k_{BG}^X(x', x'') = c(x' - X)(x'' - X)\delta(x' - x'').$$

Критерий L может быть записан виде, содержащем точку оценивания X :

$$L = (\vec{\alpha}^T S \vec{\alpha}),$$

где

$$\begin{aligned} S_{ij} &= c \int \int dx' dx'' p_i(x') p_j(x'') (x' - X)(x'' - X) \delta(x' - x'') = \\ &= c \int p_i(x) p_j(x) (x - X)^2 dx. \end{aligned}$$

Величина L связывается с разрешающей способностью метода по координате x .

Подставляя вместо линейно комбинации $(\vec{\alpha} \cdot \vec{p})$ равномерной на интервале Δ Распределение (см. рис. 10.11):

$$\begin{aligned} L &= c \int (\vec{\alpha} \cdot \vec{p})^2 (x - X)^2 dx \Rightarrow c \int_{x-\Delta/2}^{x+\Delta/2} \frac{1}{\Delta^2} (x - X)^2 dx = \\ &= \frac{c}{3\Delta^2} (x - X)^3 \Big|_{x-\Delta/2}^{x+\Delta/2} = \frac{\Delta \cdot c}{12}, \\ c = 12 &\Rightarrow L = \Delta. \end{aligned}$$

Величина L численно равна величине интервала Δ . При этом Δ интерпретиру-

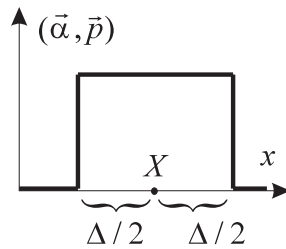


Рис. 10.11: Графическое представление функции окна.

ется как *длина разрешения* по координате x .

Вторая часть критерия W_{BG} связана с ошибкой измерения ε

$$N = (\vec{\alpha}^T K_{\varepsilon\varepsilon} \vec{\alpha}),$$

где $K_{\varepsilon\varepsilon} = E(\varepsilon\varepsilon^T)$ ковариационная матрица ошибки измерения.

Окончательно, в качестве критерия оптимального оценивания функционала l^X предлагается минимизировать сумму взвешенных квадратичных по α форм L и N

$$W_{BG} = (1 - \gamma)L + \gamma N \quad (\gamma \in [0, 1])$$

или

$$W_{BG} = (1 - \gamma) \langle l - \hat{l} | K_{BG}^X | l - \hat{l} \rangle + \gamma (\vec{\alpha}^T K_{\varepsilon\varepsilon} \vec{\alpha})$$

при условии

$$\langle \hat{l} | 1 \rangle = 1 \Rightarrow (\vec{\alpha}^T \cdot \vec{p}) = 1.$$

Для решения задачи мы будем использовать метод неопределенных множителей Лагранжа

$$\begin{aligned} F(\vec{\alpha}) &= (1 - \gamma) \vec{\alpha}^T S \vec{\alpha} + \gamma \vec{\alpha}^T K_{\varepsilon\varepsilon} \vec{\alpha} - 2\lambda (\vec{\alpha}^T \cdot \vec{p} - 1), \\ \frac{\partial F}{\partial \alpha_i} &= 0, \quad i = 1, 2, \dots, n, \\ 2[(1 - \gamma)S + \gamma K_{\varepsilon\varepsilon}] \vec{\alpha} &= 2\lambda \vec{p}. \end{aligned}$$

Введем обозначение

$$B = [(1 - \gamma)S + \gamma K_{\varepsilon\varepsilon}], \quad \hat{\alpha} = B^{-1}\vec{p}\lambda.$$

Умножим оценку $\hat{\alpha}$ на вектор \vec{p}^T

$$\vec{p}^T | \hat{\alpha} = B^{-1}\vec{p}\lambda \Rightarrow \vec{p}^T \hat{\alpha} = \vec{p}^T B^{-1}\vec{p}\lambda.$$

Используя условие

$$\vec{p}^T \vec{\alpha} = 1,$$

получаем для λ

$$1 = \vec{p}^T B^{-1}\vec{p}\lambda \Rightarrow \lambda = [\vec{p}^T B^{-1}\vec{p}]^{-1}.$$

Окончательно получаем оценку

$$\hat{\alpha} = \frac{B^{-1}\vec{p}}{[\vec{p}^T B^{-1}\vec{p}]}.$$

Проведем сравнение критерия оптимальности W_{BG} критерия реконструктивной томографии W_{RT} , который запишем в виде

$$W_{RT} = \langle l_{RT}^X - \hat{l} | K_{RT}^X | l_{RT}^X - \hat{l} \rangle + (\vec{\alpha}^T K_{\varepsilon\varepsilon} \vec{\alpha}).$$

Априорная информация относительно $\nu(\delta\theta)$ дается оператором K_{RT} , который действительно является ковариационным оператором в случае нормального распределения

$$K_{RT}^X = K_{\nu\nu}.$$

Легко показать, что в случае W_{BG} первая часть представляет собой квадратичную форму смещения (по аналогии с W_{RT}), оператор K_{BG} отвечает за “априорную информацию”. Но оператор K_{BG} имеет сингулярное ядро, которое зависит от оцениваемой точки X . Так как выбор оцениваемой точки произволен, то “априорные представления” в BG -методе волевым образом трансформируются вместе с изменением точки X , что недопустимо при обработке фиксированного объема экспериментальных данных.

Линейное условие $\langle \hat{l} | 1 \rangle$, используемое в W_{BG} , привнесено искусственно и обусловлено двумя факторами:

1. Оцениваемый функционал сосредоточен в точке X .
2. Ядро K_{BG}^X обращается в нуль в этой же точке.

Алгоритм реконструктивной томографии W_{RT} свободен от ограничений характерных для критерия W_{BG} . Критерий W_{RT} в полной мере учитывает как априорные представления об упругих полях, так и статистическую структуру ошибок наблюдений. На рис. 10.12, а представлено ядро критерия K_{BG} . Ядро K_{RT} в случае некоррелированного процесса и марковского процесса представлено на рис. 10.12, б и с соответственно. Это представление характеризует задание априорной информации в BG - и RT -методах (см. рис. 10.12).

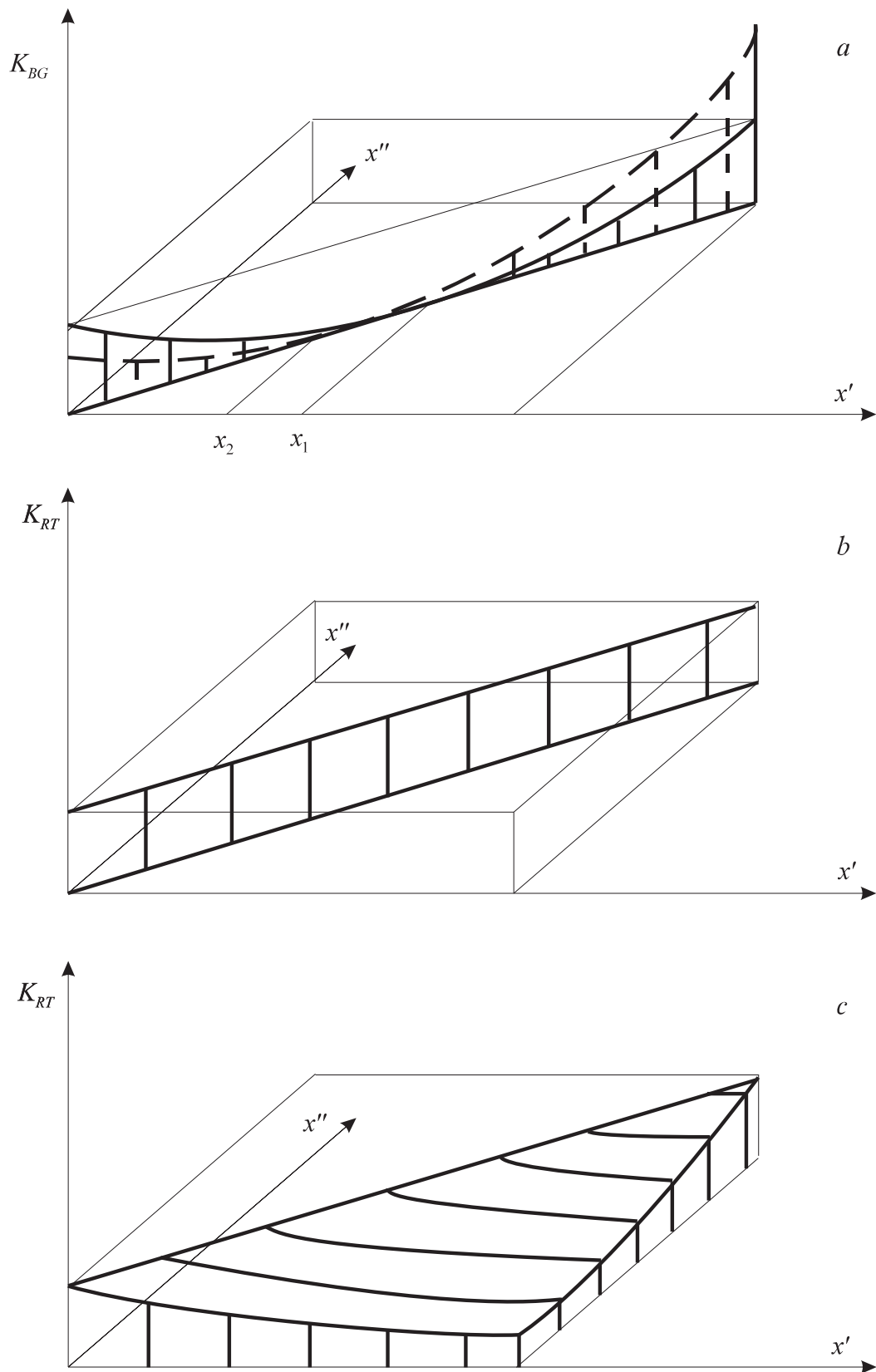


Рис. 10.12: Ядра K_{BG} (а) и K_{RT} в случаях некоррелированного процесса (b) и марковского процесса (с).

Глава 11

Многофакторный анализ временных рядов

Рассмотрим временной ряд

$$\{x_i\}_{i=1}^N \Rightarrow x_i = f((i-1)\Delta t), \quad i = 1, 2, \dots, N.$$

Функция $f(t)$ может быть случайной.

I. Преобразование одномерного временного ряда в многомерный временной ряд

Выберем число $M < N$, которое назовем *максимальным числом главных компонент (МЧГК)*. Представим первые M элементов временного ряда x_i как первую строку матрицы X . В качестве элементов второй строки возьмем элементы временного ряда от x_2 по x_{M+1} включительно. Последнюю строку с номером $k = N - M + 1$ составят M элементов временного ряда x_k, x_{k+1}, \dots, x_N :

$$X = (x_{ij})_{i,j=1}^{k,M} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_M \\ x_2 & x_3 & x_4 & \dots & x_{M+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_k & x_{k+1} & x_{k+2} & \dots & x_N \end{bmatrix}. \quad (11.1)$$

Подобную матрицу с элементами $x_{ij} = x_{i+j-1}$, можно рассматривать как M -размерную выборку объема k .

II. Анализ главных компонент: сингулярное разложение выборочной корреляционной матрицы

1. Вычисление средних значений и стандартных отклонений для столбцов матрицы X :

$$\bar{x}_j = \frac{1}{k} \sum_{i=1}^k x_{i+j-1}, \quad s_j = \sqrt{\frac{1}{k} \sum_{i=1}^k (x_{i+j-1} - \bar{x}_j)^2}, \quad (11.2)$$

где \bar{x}_j является скользящим средним, а s_j — стандартным отклонением с окном шириной k .

2. Введем центрированную матрицу X^* :

$$x_{ij}^* = \frac{(x_{ij} - \bar{x}_j)}{s_j}, \quad i = 1, \dots, k; \quad j = 1, \dots, M. \quad (11.3)$$

3. Вычислим выборочную корреляционную матрицу

$$R = \frac{1}{k} X^{*T} X^* \quad (11.4)$$

с элементами

$$r_{ij} = \frac{1}{k} \sum_{l=1}^k \frac{1}{s_i s_j} (x_{i+l-1} - \bar{x}_i)(x_{j+l-1} - \bar{x}_j).$$

4. Вычислим собственные числа и собственные векторы матрицы R (т.е. выполним сингулярное разложение):

$$R = P \Lambda P^T, \quad (11.5)$$

где

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_M \end{bmatrix}$$

— диагональная матрица собственных чисел и

$$P = (p_1, p_2, \dots, p_M) = \begin{bmatrix} p_{11} & p_{21} & \dots & p_{M1} \\ p_{12} & p_{22} & \dots & p_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1M} & p_{2M} & \dots & p_{MM} \end{bmatrix}$$

— ортогональная матрица собственных векторов матрицы R .

В этом случае выполняются следующие равенства:

$$P^T = P^{-1}, \quad P^T P = P P^T = I_M, \\ \Lambda = P^T R P, \quad \sum_{i=1}^M \lambda_i = M, \quad \prod_{i=1}^M \lambda_i = \det R.$$

Интерпретация матриц P и Λ . Матрицу P можно рассматривать как матрицу перехода от матрицы X^* к матрице, составленной из главных компонент

$$X^* P = Y = (y_1, y_2, \dots, y_M).$$

Если рассматривать случайный временной ряд x_i , то собственные числа матрицы R — выборочные дисперсии главных компонент, а их квадратные корни —

среднеквадратичные отклонения. Выборочные среднеквадратичные отклонения пропорциональны полуосям эллипсоидов, задаваемых матрицей R .

Если временной ряд $\{x_i\}$ является стационарным, в том смысле, что любая часть временного ряда длины $s = \min(M, k)$ несет в себе всю информацию относительно процесса $\{x_i\}$, то можно усреднить матрицу R вдоль диагоналей и получать автокорреляционную функцию $r(\tau)$:

$$r(\tau) = \frac{1}{M - \tau} \sum_{j=1}^{M-\tau} r_{j, j+\tau}, \quad \tau = 1, \dots, M. \quad (11.6)$$

III. Отбор главных компонент

Как было упомянуто выше, главные компоненты определяются следующим образом:

$$Y = X^*P = (y_1, y_2, \dots, y_M).$$

Нормируем главные компоненты (при $\lambda_M \neq 0$)

$$Y^* = Y\Lambda^{-1/2} = (y_1^*, y_2^*, \dots, y_M^*).$$

В этом случае главные компоненты являются ортонормированными:

$$Y^{*T}Y^* = I_M.$$

Таким образом, получено разложение исходного M -мерного процесса на ортогональные составляющие.

Преобразование

$$y_j = X^*p_j$$

очень близко к линейному преобразованию исходных временных рядов дискретным оператором свертки

$$\begin{aligned} y_j(l) &= \sum_{q=1}^M X_{lq}^* p_{jq} = \sum_{q=1}^M (x_{l+q-1} - \bar{x}_q) \frac{p_{jq}}{s_q} = \\ &= \sum_{q=1}^M x_{l+q-1} \frac{p_{jq}}{s_q} - \sum_{q=1}^M \bar{x}_q \frac{p_{jq}}{s_q}. \end{aligned}$$

Иначе говоря, процедура многофакторного анализа формирует набор линейных фильтров в которых собственные векторы играют роль переходных функций.

Визуальный анализ собственных векторов и главных компонент, полученных после линейной фильтрации, может дать полезную информацию относительно структуры процесса.

Может оказаться полезной информация, полученная в результате наблюдения за двумерными диаграммами (по x -оси — p_i или y_i , по y -оси p_j или y_j).

Пример. Если двумерная диаграмма для p_i и p_{i+1} близка к окружности, то собственные векторы можно интерпретировать как пару \sin - \cos на языке преобразования Фурье. Рассмотрим функции

$$\begin{cases} x(t) = a(t) \cos(\omega(t)t + \psi(t)), \\ y(t) = a(t) \sin(\omega(t)t + \psi(t)), \end{cases}$$

полагая $a(t)$, $\omega(t)$ и $\psi(t)$ медленно изменяющимися функциями

$$\begin{cases} x(t) = r(t) \cos(\alpha(t)t), \\ y(t) = r(t) \sin(\alpha(t)t), \end{cases}$$

где

$$\begin{aligned} r(t) &= \sqrt{x^2(t) + y^2(t)} = a(t), \\ \alpha(t) &= \arctan \frac{y(t)}{x(t)} = \omega(t)t + \psi(t). \end{aligned}$$

Выберем две последовательности с задержкой τ и найдем разность полярных углов:

$$\begin{aligned} \Delta\alpha &= \alpha(t + \tau) - \alpha(t) = \omega(t + \tau)(t + \tau) + \\ &+ \psi(t + \tau) - \omega(t)t - \psi(t) = \\ &= (\omega(t + \tau) - \omega(t))t + \psi(t + \tau) - \psi(t) = \\ &= (\omega'(t)\tau + o(\tau))t + (\omega(t) + \\ &+ \omega'(t)\tau + o(\tau))\tau + \psi'(t)\tau + o(\tau). \end{aligned}$$

Если $\omega(t)$ и $\psi(t)$ — медленно изменяющиеся функции, такие что $\omega'(t)$ и $\psi'(t)$ имеют порядок малости $O(\tau^2)$, можно записать

$$\begin{aligned} \Delta\alpha(t) &= \omega(t)\tau = \frac{2\pi\tau}{T(t)}, \\ \omega(t) &= \frac{\Delta\alpha(t)}{\tau} \quad \text{и} \quad T(t) \approx \frac{2\pi\tau}{\Delta\alpha(t)}. \end{aligned}$$

Таким образом, вычисляя разность полярных углов, мы можем оценить мгновенную частоту $\omega(t)$ и мгновенный период $T(\tau)$ “гармонических колебаний” отвечающих выбранной паре собственных векторов.

Следует отметить, что невозможно оценить частоты, превышающие

$$\omega_N = \frac{\pi}{\tau}.$$

IV. Восстановление одномерных временных рядов

С учетом ортогональности матрицы P можно записать

$$X^* = YP^T = (y_1, y_2, \dots, y_M) \cdot \begin{pmatrix} p_1^T \\ p_2^T \\ \vdots \\ p_M^T \end{pmatrix} =$$

$$= \sum_{l=1}^M y_l P_l^T = \sum_{l=1}^M X_l^*.$$

Снятие нормировки и центрирование матрицы X^*

$$X = \bar{x} I_k^T + X^* S = X_0^* + \sum_{l=1}^M X_l^* S = \sum_{l=0}^M X_l^* S.$$

В результате получаем исходную матрицу как сумму $M + 1$ матриц.

Преобразование к начальному временному ряду может быть выполнено усреднением по диагоналям, прилегающим к главной диагонали. Определим оператор усреднения \mathcal{A} :

$$x = \mathcal{A}(X) = \sum_{l=0}^M \mathcal{A}(X_l^* S). \quad (11.7)$$

Таким образом, получено представление начального временного ряда как суммы $M + 1$ временных рядов.

Следует отметить, что при применении алгоритма, интерактивный подход к исследованию (т. е. “диалог” между исследователем и ЭВМ) является естественной его особенностью.

Выбор параметров

Основной входной параметр M алгоритма МЧГК следует полагать меньшим, чем $N/2$.

Выбор величины M зависит от типа решаемой задачи:

1. *Наблюдение латентной периодичности.*

Вычисляем собственные числа при максимальном значении M и определяем число l собственных чисел, удовлетворяющих условию $\lambda_i > 0$. На следующем шаге анализ выполняется при $M = l$.

2. *Сглаживание временных рядов.*

В этом случае мы рассматриваем алгоритм как алгоритм фильтрации. При фильтрации следует уменьшать число главных компонент, участвующих в восстановлении временного ряда, при этом большим значениям M , соответствует более узкая полоса фильтрации.

3. *Выделение периодичности с известным периодом.*

M должно быть равно периоду T и N должно быть кратно периоду T .

Следует отметить, что рассматриваемый алгоритм очень стабилен к вариациям величины M .

Промежуточные результаты интерпретации.

1. Собственные числа корреляционной матрицы, отвечающей M -размерному представлению временного ряда.

2. Собственные векторы корреляционной матрицы.
3. Главные компоненты, отвечающие M -размерному представлению временного ряда.
4. Восстановленный временной ряд при различном числе главных компонент.

Отметим два предельных случая.

- $M \ll N$. В этом случае можно рассматривать собственные векторы как переходные функции линейного фильтра, а главные компоненты — как результат действия такого фильтра.
- $M \simeq N/2$. При таком соотношении M и N алгоритм можно рассматривать как аппроксимационный и применять, например, для выделения гармонических составляющих временного ряда.

Приложение А

КОМПЬЮТЕРНЫЕ УПРАЖНЕНИЯ

Компьютерные упражнения, представленные в Приложении, составляют основу вычислительного практикума по курсам лекций “Статистические методы обработки геофизических данных” и “Обратные задачи геофизики”, читаемых магистрантам и аспирантам кафедры физики Земли на физическом факультете Санкт-Петербургского государственного университета в течение последних десяти лет. Выполнение компьютерных упражнений позволит читателю глубже понять содержание и возможности представленных в книге методов анализа и обработки геофизической информации.

Основой компьютерных упражнений является набор программ¹, написанных на языке пакета МАТЛАБ. Программы содержат краткое описание входных и выходных данных, демонстрационный пример и рекомендации по изменению входных данных². Здесь не приводятся руководства по работе в пакете МАТЛАБ, поскольку для выполнения компьютерных упражнений практически не требуется (по крайней мере на начальном этапе) знания языка пакета МАТЛАБ.

А.1 Статистические методы

А.1.1 Численное моделирование случайных величин

В пакете МАТЛАБ имеются встроенные функции, реализующие генераторы случайных чисел с равномерным распределением на интервале $(0, 1)$ — `rand` и с нормальным распределением $(N(0, 1))$ — `randn`. Следующая программа (скрипт

¹ Тексты программ, написанные на языке МАТЛАБ, находятся на FTP сервере Лаборатории динамики упругих сред физического факультета С.-Петербургского государственного университета (`ldus.phys.spbu.ru`).

² Тексты программ на языке МАТЛАБ и демонстрационные примеры, представленные в Приложении, были подготовлены Ю.В. Киселевым в Институте геофизики Лозанского университета. Авторы считают своим приятным долгом выразить признательность профессору R.Olivier за создание превосходных условий для работы и дружескую атмосферу.

`r0a.m`), в зависимости от значения переменной `indd` генерирует случайные выборки и реализует их графическое представление.

```
% Script r0a.m
% Генераторы случайных чисел
% indd=1 : равномерное распределение в интервале (0,1)
% indd=2 : нормальное распределение (N(0,1))
clear
% ВВОД ДАННЫХ
indd=1;
m=25;      % объем выборки
% КОНЕЦ ВВОДА
n=1:m;     % n=1,2,...,m
if indd==1
    s=rand(1,m);    % s : вектор-строка
else
    s=randn(length(n),1);    % s : вектор-столбец; length(n)=m
    s=s'           % s : вектор-строка
end
rect=[0.1 0.1 .8 0.5];    % размер рисунка
axes('position',rect)
stem(n,s);    % графическое представление
```

Упражнения

1. С использованием скрипта `r0a.m` получить графическое представление выборок из равномерного и нормального распределений для различных объемов выборок.

2. Выполнить задание из предыдущего упражнения, используя командную строку.

А.1.2 Построение гистограмм

Гистограмма характеризует вероятностные свойства случайной величины по выборке конечных размеров. Скрипт `p0hist.m` реализует построение нормированной гистограммы по выборке из нормального распределения $N(0, 1)$. Графический вывод содержит как гистограмму, так и плотность вероятностей соответствующего ей нормального распределения.

```
%Script p0hist.m
% Построение гистограммы
% nr : объем выборки
clear
% ВВОД ДАННЫХ
nr=500;    % объем выборки
```

```

del=0.5;      % размер интервала для разбиения оси x
% КОНЕЦ ВВОДА
x=-5:del:5; % создание интервалов на оси x
d=randn(1,nr); % d : выборка из нормального
               %распределения (вектор-строка)
% hist(d,x);
[n,x]=hist(d,x);
bar(x,n./(nr.*del)); % графическое представление
                    % нормированной гистограммы
hold on
d2=d_gauss(x,0,1); % нормальная плотность вероятностей
                  % (вектор-строка)
plot(x,d2,'g') % графическое представление нормальной
               % плотности вероятностей

```

Упражнения

1. Сравнить гистограммы, построенные по выборкам различного объема (например 100 – 500).

2. Сравнить гистограммы, построенные для различного размера интервалов разбиения оси x (например 4 – 20), без изменения минимальной и максимальной точек на оси x , задающих область определения гистограммы.

A.1.3 Описание случайных переменных

Ниже приведен ряд распределений дискретных случайных величин и плотностей вероятностей непрерывных случайных величин.

Бета-распределение

Плотность вероятностей

$$f_{\xi}(x|\alpha, \beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, & \text{если } 0 < x < 1, \\ 0, & \text{если } x \leq 0 \text{ или } x \geq 1, \end{cases} \quad (\text{A.1})$$

где $\alpha > 0$, $\beta > 0$;

$$M\xi = \frac{\alpha}{(\alpha + \beta)}, \quad D\xi = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Графическое представление (A.1) реализует скрипт `d0beta.m`. Функция `d=d_beta(x, a1, bt)` возвращает вектор-строку `d` — значения плотности вероятностей бета-распределения.

Биномиальное распределение

Графическое представление плотности вероятностей реализует скрипт `d0bin.m` (см. 1.6.1). Функция `d=d_bin(n,p)` возвращает вектор-строку `d` — значения вероятностей биномиального распределения (см. 1.6.1).

Ниже приводится текст скрипта `d0bin.m`.

```
% Script d0bin.m
% биномиальное распределение вероятностей
% call function d=d_bin(n,p)
% n : n=1,2,...; - максимальная величина аргумента x
%     x=0,1,...,n
% p : 0 < p < 1
% d : вектор-строка, содержит распределение вероятностей
%
clear
% ВВОД ДАННЫХ
n=50;
p=0.7;
% КОНЕЦ ВВОДА
d=d_bin(n,p);
x=0:n;
% графическое представление
rect=[0.1 0.1 .8 0.5]; % размер рисунка
axes('position',rect)
stem(x,d);
comp_prob=sum(d) % вычисление суммы вероятностей
math_exp= n .* p % математическое ожидание
stand_dev=sqrt(n .* p .* (1-p)) % стандартное отклонение
m1=sum(x.*d) % математическое ожидание
s1=sqrt(sum((x-m1).^2.*d)) % стандартное отклонение
pause
% сравнение с нормальным распределением
hold on
dn=d_gauss(x,math_exp,stand_dev);
plot(x,dn,'g');
```

В процессе работы программы на экран выводятся значения математического ожидания и стандартного отклонения, вычисленные как по аналитическим выражениям, так и полученные суммированием взвешенных значений вероятностей (“по определению”).

Ниже приводится текст функции `d=d_bin(n,p)`.

```
function d=d_bin(n,p);
% биномиальное распределение d=d_bin(n,p)
% n : n=1,2,...; - максимальное значение аргумента x
```

```

%      x=0,1,...,n
% p : 0 < p < 1
%
n1=prod(1:n);
d=zeros(1,n+1);      % задание вектора-строки d нулевыми
                    % значениями
for m=0:n
    if m==0
        m1=1;
    else
        m1=prod(1:m);
    end
    nm=n-m;
    if nm==0
        nm1=1;
    else
        nm1=prod(1:nm);
    end
    d(m+1)=(n1/(m1*nm1))*p^m * (1-p)^nm;
end

```

Распределение Коши

Графическое представление плотности вероятностей распределения Коши (см. 1.7.9). реализует скрипт `d0cauch.m`. Функция `d=d_cauch(x)` возвращает вектор-строку `d` — значения плотности вероятностей распределения Коши.

χ^2 -распределение

Графическое представление плотности вероятностей χ^2 -распределения реализует скрипт `d0chisq.m` (см. 1.7.4). Функция `d=d_chisq(x,n)` возвращает вектор-строку `d` — значения плотности вероятностей χ^2 -распределения.

Ниже приводится текст скрипта `d0chisq.m`.

```

% Script d0chisq.m
% плотность вероятностей для  $\chi^2$ -распределения
% call d=d_chisq(x,n)
% n : число степеней свободы
% x : значения случайной переменной, a < x < b
% d : вектор-строка, возвращает значения плотности
      % вероятностей
clear
% ВВОД ДАННЫХ
n=50
% задание интервала на оси x для вычисления плотности
      % вероятностей

```

```

a=0;          % (a < x < b)
b=100.0;
nx=500;      % объем выборки
% КОНЕЦ ВВОДА
del=(b-a)/(nx-1);      % задание значений аргумента
x=a:del:b;
d=d_chisq(x,n);      % вектор-строка
plot(x,d);          % графическое представление
comp_prob=trapz(x,d) % вычисление полной вероятности
math_exp=n          % математическое ожидание
stand_dev=sqrt(2 .* n) % стандартное отклонение
m1=trapz(x,x.*d)    % математическое ожидание
s1=sqrt(trapz(x,(x-m1).^2.*d)) % стандартное отклонение
%
pause
% сравнение с нормальным распределением
hold on
dn=d_gauss(x,math_exp,stand_dev);
plot(x,dn,'g');     % графическое представление
                    % нормального распределения

```

В процессе работы программы на экран выводятся значения математического ожидания и стандартного отклонения, вычисленные как по аналитическим выражениям, так и полученные интегрированием взвешенных значений вероятностей (“по определению”). Ниже приводится текст функции `d=d_chisq(x,n)`.

```

function d=d_chisq(x,n)
%  $\chi^2$ -распределение d=d_chisq(x,n)
% n : число степеней свободы
% x : вектор-строка значений аргумента(0 < x < inf)
% d : вектор-строка -- значения плотности вероятностей
%
d=((x./2).^((n./2)-1).*exp(-x ./ 2))/(2.*gamma(n./2));

```

Экспоненциальное распределение

Графическое представление экспоненциальной плотности вероятностей реализует скрипт `d0expon.m` (см. 1.7.7). Функция `d=d_expon(x,1)` возвращает вектор-строку `d` — значения плотности вероятностей экспоненциального распределения.

Распределение Фишера

Графическое представление плотности вероятностей распределения Фишера реализует скрипт `d0fish.m` (см. 1.7.6). Функция `d=d_fish(x,n1,n2)` возвращает вектор-строку `d` — значения плотности вероятностей распределения Фишера.

Гамма распределение

Плотность вероятностей гамма-распределения дается формулой

$$f_{\xi}(x|\alpha, \beta) = \begin{cases} \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & \text{если } x \geq 0, \\ 0, & \text{если } x < 0, \end{cases} \quad (\text{A.2})$$

где $\alpha > 0$, $\beta > 0$;

$$M\xi = \frac{\alpha}{\beta}, \quad D\xi = \frac{\alpha}{\beta^2}.$$

Графическое представление плотности вероятностей гамма-распределения реализует скрипт `d0gam.m`. Функция `d=d_gam(x, al, bt)` возвращает вектор-строку `d` — значения плотности вероятностей гамма-распределения.

Нормальное (одномерное) распределение

Графическое представление плотности вероятностей нормального распределения реализует скрипт `d0gauss.m` (см. 1.7.1). Функция `d=d_gauss(x, xm, s)` возвращает вектор-строку `d` — значения плотности вероятностей нормального распределения.

Геометрическое распределение

Графическое представление геометрического распределения реализует скрипт `d0geom.m` (см. 1.6.3). Функция `d=d_geom(n, p)` возвращает вектор-строку `d` — значения вероятностей геометрического распределения.

Гипергеометрическое распределение

Распределение вероятностей гипергеометрического распределения дается формулой

$$f_{\xi}(x|n, a, b) = \begin{cases} \frac{C_x^a C_{n-x}^b}{C_n^{a+b}}, & x = 0, 1, 2, \dots, n, \\ 0, & x \neq 0, 1, 2, \dots, n, \end{cases} \quad (\text{A.3})$$

где $a = 0, 1, 2, \dots$, $b = 0, 1, 2, \dots$, $n = 0, 1, 2, \dots$, $a + b \geq n$,

$$M\xi = \frac{nab}{(a+b)^2}, \quad D\xi = \frac{a+b-n}{a+b-1}.$$

Графическое представление гипергеометрического распределения реализует скрипт `d0hurg.m`. Функция `d=d_hurg(n, a, b)` возвращает вектор-строку `d` — значения вероятностей гипергеометрического распределения.

Распределение Лапласа

Графическое представление плотности вероятностей распределения Лапласа реализует скрипт `d0lapl.m` (см. 1.7.8). Функция `d=d_lapl(x, l)` возвращает вектор-строку `d` — значения плотности вероятностей распределения Лапласа.

Логарифмически нормальное распределение

Графическое представление плотности вероятностей логарифмически нормального распределения реализует скрипт `d0lnorm.m` (см. 1.7.10). Функция `d=d_lnorm(x, xm, s)` возвращает вектор-строку `d` — значения плотности вероятностей логарифмически нормального распределения.

Распределение Пуассона

Графическое представление распределения Пуассона реализует скрипт `d0poiss.m` (см. 1.6.2). Функция `d=d_poiss(n, a)` возвращает вектор-строку `d` — значения вероятностей распределения Пуассона.

Распределение Стьюдента

Графическое представление плотности вероятностей распределения Стьюдента реализует скрипт `d0stu.m` (см. 1.7.5). Функция `d=d_stu(x, n)` возвращает вектор-строку `d` — значения плотности вероятностей распределения Стьюдента.

Равномерное распределение

Графическое представление плотности вероятностей равномерного распределения реализует скрипт `d0unif.m` (см. 1.7.3). Функция `d=d_unif(x, al, bt)` возвращает вектор-строку `d` — значения плотности вероятностей равномерного распределения.

Двумерное нормальное распределение

Графическое представление плотности вероятностей двумерного нормального распределения (см. 1.7.2) реализует скрипт `d0gauss2.m`. Функция `d=d_gauss2.m` возвращает матрицу `d` — значения плотности вероятностей двумерного нормального распределения.

Упражнения

1. Сравнить биномиальное распределение с плотностью вероятностей нормального распределения для различных значений параметров n ($n = 1, 5, 20$) и p ($p = 0.1, 0.5, 0.9$).

2. Сравнить распределение Пуассона с плотностью вероятностей нормального распределения для различных значений параметра a ($a = 0.1, 0.5, 1.0, 5.0$).

3. Сравнить плотность вероятностей χ^2 -распределения с плотностью вероятностей нормального распределения при различных значениях параметра n ($n = 2, 5, 10, 20$).

4. Сравнить плотность вероятностей распределения Стьюдента с плотностью вероятностей нормального распределения при различных значениях параметра n ($n = 2, 5, 10, 20$).

5. Выполнить графическое построение двумерной нормальной плотности при различных значениях математических ожиданий m_1 и m_2 ($m_1 = 0, 1, 10$; $m_2 = 0, 1, 10$), стандартных отклонений σ_1 и σ_2 ($\sigma_1 = 0, 1, 10$; $\sigma_2 = 0, 1, 10$) и коэффициента корреляции ρ ($\rho = -0,95; 0; 0,95$).

6. Выполнить графическое представление других распределений вероятностей, упомянутых в п. А.1.3.

А.1.4 Генераторы случайных чисел

Существует широкий спектр алгоритмов генерирования случайных чисел на ЭВМ для различных типов распределений. Ниже приведены простейшие алгоритмы для генерации случайных чисел, реализованные в скриптах `rnd0ald.m` и `rnd0alc.m` соответственно для дискретных и непрерывных случайных величин, упомянутых в п. А.1.3.

При этом в обоих случаях для создания выборок случайных чисел, удовлетворяющих различным распределениям, используется генератор случайных чисел с равномерным распределением на интервале $[0, 1]$, т. е. с плотностью вероятностей вида

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1, \\ 0, & x < 0 \text{ или } x > 0. \end{cases} \quad (\text{А.4})$$

Метод моделирования значений дискретных случайных величин ξ с распределением вероятностей $p_m = P(\xi = x_m)$ основан на следующем равенстве:

$$P\left(\sum_{n=0}^{m-1} p_n \leq \alpha < \sum_{n=0}^m p_n\right) = p_m, \quad (\text{А.5})$$

где величина α принадлежит равномерному распределению (А.4).

В случае непрерывных случайных величин, задаваемых функцией распределения $F(x)$, генератор случайных чисел построен с использованием выражения

$$\xi = F^{-1}(\alpha), \quad (\text{А.6})$$

где $F^{-1}(\alpha)$ — обратная функция распределения, α принадлежит равномерному распределению (А.4), ξ принадлежит распределению с функцией распределения $F(x)$. Отметим, что если $F(x)$ — строго монотонная непрерывная функция, то справедливы следующие равенства:

$$P(F^{-1}(\alpha) < x) = P(\alpha < F(x)) = F(x). \quad (\text{А.7})$$

Генераторы случайных величин, подчиняющиеся равномерному и нормальному распределениям с произвольными параметрами, реализованы соответственно в скриптах `rnd0uni.m` и `rnd0gau.m`. Скрипт `rnd0n2.m` реализует генератор двумерного нормального распределения.

Критерий согласия Колмогорова

В скрипте `g0kolm.m` реализован критерий согласия Колмогорова о принадлежности выборки случайных величин (эмпирической функции распределения) заданному распределению (распределению, заданному своей функцией распределения).

Критерий Колмогорова дает вероятность события, заключающегося в том, что максимальная верхняя грань модуля разности функции распределения $F_\xi(x)$ и ее выборочного аналога $F_N^*(x)$ меньше, чем величина γ/\sqrt{N} ,

$$\begin{aligned} \lim_{N \rightarrow \infty} P \left(\sup_{-\infty < x < \infty} |F_\xi(x) - F_N^*(x)| < \gamma/\sqrt{N} \right) = \\ = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2\gamma^2), \end{aligned} \quad (\text{A.8})$$

где N — объем выборки. Критерий реализован для непрерывных распределений из п. А.1.3.

Упражнения

1. Используя скрипты `rnd0alc.m` и `rnd0ald.m` из п. А.1.4, сравнить величины математического ожидания и стандартного отклонения с их выборочными аналогами для различных объемов выборок (например 10 – 500).

2. Используя скрипт `rnd0n2.m`, сравнить значения параметров двумерного нормального распределения (математических ожиданий, стандартных отклонений, коэффициента корреляции) с их выборочными аналогами для различных объемов выборки (например 100 – 5000).

3. Используя скрипт `g0kolm.m`, вычислить вероятность, отвечающую критерию Колмогорова, для некоторых типов распределений (нормального, χ^2 , Стьюдента). Обратит внимание на поведение выборочной функции распределения при различных объемах выборки (например 10 – 1000).

А.1.5 Доверительные интервалы

Скрипт `g0conf.m` реализует построение доверительного интервала для выборочного среднего (см. п. 1.7.12) случайной величины, распределенной по нормальному закону.

Упражнения

1. Используя скрипт `g0conf.m`, найти доверительные интервалы для выборочного среднего случайной величины, распределенной по нормальному закону при различных значениях доверительной вероятности β (например $\beta=0.2, 0.5, 0.9$) и объема выборки (например 10 – 100).

А.1.6 Временные ряды

Коррелированные и некоррелированные процессы

Скрипт `r0covgr.m` генерирует некоррелированный временной ряд и цепь Маркова, выполняет построение выборочной ковариационной матрицы.

Автокорреляционная и взаимокорреляционная функции

Скрипт `r0corf.m` выполняет построение автокорреляционных и взаимокорреляционных функций для некоррелированных временных рядов и цепей Маркова.

Упражнения

1. Используя скрипт `r0covgr.m`, вычислить выборочные ковариационную и корреляционную матрицы для некоррелированных и коррелированных временных рядов различной длины (например 100, 1000, 10000).

2. Используя скрипт `r0corf.m`, сравнить автокорреляционные и взаимокорреляционные функции для коррелированных и некоррелированных временных рядов.

А.2 Преобразования

А.2.1 Преобразование Фурье

Функции МАТЛАБ `fft` и `ifft` реализуют дискретное преобразование Фурье и обратное дискретное преобразование Фурье соответственно, которые задаются формулами вида

$$y_k = \sum_{r=0}^{N-1} x_r \exp \left[-i \frac{2\pi k r}{N} \right], \quad k = 0, \dots, N-1 \quad (\text{A.9})$$

(преобразование Фурье),

$$x_r = \frac{1}{N} \sum_{k=0}^{N-1} y_k \exp \left[i \frac{2\pi k r}{N} \right], \quad r = 0, \dots, N-1 \quad (\text{A.10})$$

(обратное преобразование Фурье).

А.2.2 Простейшие сигналы и их спектральные характеристики

Скрипт `r0wlet.m` создает модель простейшей сейсмической трассы: сигнал и некоррелированный шум с нормальным распределением. Вычисляются спектральные характеристики трассы, автокорреляционная функция и спектр мощности

(см. 3.5, 3.6, 3.7). Функция `p_wlet.m`, вызываемая скриптом `p0wlet.m`, создает различные сигналы.

1. Дискретная дельта-функция:

$$x_r = \begin{cases} 1, & \text{если } r = j, \\ 0, & \text{если } r \neq j. \end{cases} \quad (\text{A.11})$$

2. Дельта-подобный сигнал:

$$x_r = \exp[-\pi(r - t_p)/t_c]^2, \quad r = 0, \dots, N - 1, \quad (\text{A.12})$$

где t_c — ширина импульса, а t_p определяет сдвиг относительно нулевого отсчета $r = 0$.

3. Импульс Риккера:

$$x_r = \exp[-(\pi * (r - (t_c + t_p))/t_c)^2] \times \\ \times (1 - 2(\pi(r - (t_c + t_p))/t_c)^2), \\ r = 0, \dots, N - 1, \quad (\text{A.13})$$

где t_c — ширина импульса, а t_p определяет сдвиг относительно нулевого отсчета $r = 0$. Пример графического вывода скрипта `p0wlet.m` дан на рис. А.1.

Упражнение

1. Используя скрипт `p0wlet.m`, вычислить спектры импульсов сигналов из формул (А.11)–(А.13) для различных значений t_p — временного сдвига и различных соотношений шум/сигнал.

А.2.3 Многофакторный анализ

Скрипт `p0cater1.m` реализует алгоритм многофакторного анализа (см. § 11).

Ниже приведен пример задания исходных данных для скрипта `p0cater1.m` при анализе простейшей записи “сигнал и шум” (рис. А.2):

```
pr=500;           % длина записи (число отсчетов)
jm=20;           % максимальное число главных
                 % компонент
unzer=2;         % число главных компонент
                 % (ненулевые собств., числа) при
                 % восстановлении
ор=3;            % задание формы сигнала
tr=100;         % сдвиг сигнала на записи
                 % (число отсчетов)
%
```

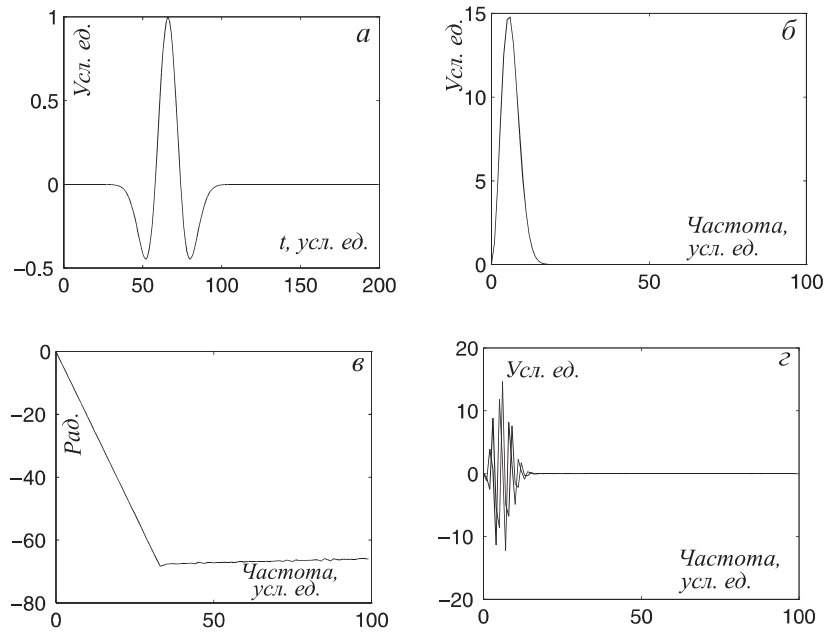


Рис. А.1: Пример графического вывода скрипта `powlet.m`. Импульс Риккера (а) и модуль его спектра (б), фаза спектра (в), вещественная и мнимая части спектра (г). По оси абсцисс отложены номера отсчетов — см. формулы (А.9), (А.10).

```
xm=0;           % математическое ожидание для шума
s=0.05;        % стандартное отклонение для шума
```

Ковариационная матрица записи и ее собственные числа (и собственные векторы) представлены на рис. А.3 и А.4 соответственно. Диаграммное представление пар главных компонент и полный набор главных компонент даны соответственно на рис. А.5 и А.6 (см. пример из § 11). Результат восстановления исходной записи с суммированием различного числа (1 – 4) главных компонент представлен на рис. А.7.

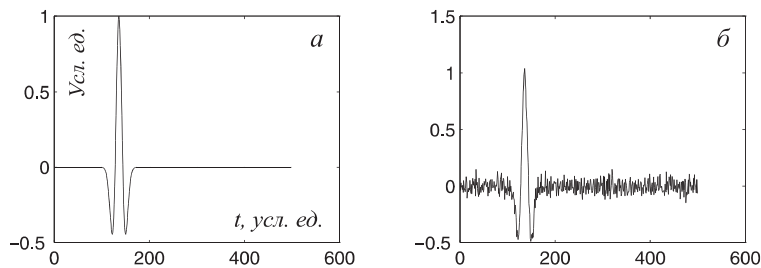
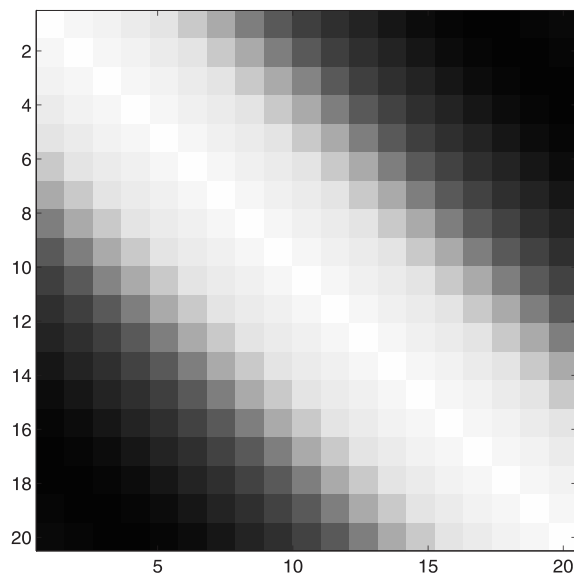


Рис. А.2: Исходная запись без шума (а) и с аддитивным некоррелированным шумом (б).

Рис. А.3: Корреляционная матрица $M = 20$.*Упражнение*

1. Используя скрипт `р0сater1.m`, выполнить многофакторный анализ записи (сигнал и нормальная шумовая компонента) для различных значений максимального числа главных компонент (например 5 – 20), для различного числа главных компонент (ненулевых собственных чисел) (например 2 – 10), при различном отношении шум/сигнал.

А.2.4 Кепстральное преобразование

Скрипт `р0сер1.m` реализует кепстральное преобразование.

Кепстральное преобразование в скрипте `р0сер1.m` реализовано следующим образом:

```
d=fft(trace);           % преобразование Фурье
                        % вектора "trace"
dabs=abs(d);           % вычисление абсолютной
                        % величины спектра
dang=unwrap(angle(d)); % вычисление фазы
dсер=fft(log(dabs)+i*dang); % преобразование Фурье
                        % логарифма спектра (кепстр)
```

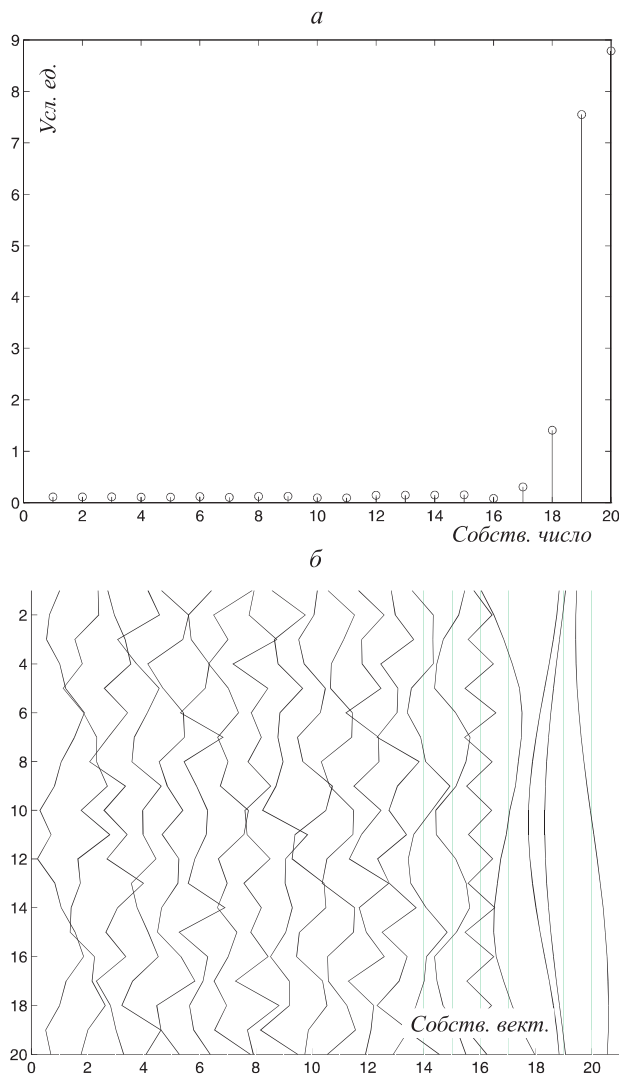


Рис. А.4: Собственные числа (а) и собственные векторы (б) корреляционной матрицы.

Обратное кепстральное преобразование реализовано следующим образом:

```

dback=ifft(dсер);           % обратное преобразование
dexp=exp(dback);           % Фурье
dback0=ifft(dexp);         % обратное преобразование
                           % Фурье
    
```

На рис. А.8–А.9 представлены два сигнала и их кепстральные характеристики соответственно. Результат свертки первого и второго сигналов сравнивается с обратным кепстральным преобразованием суммы их кепстров — рис. А.10.

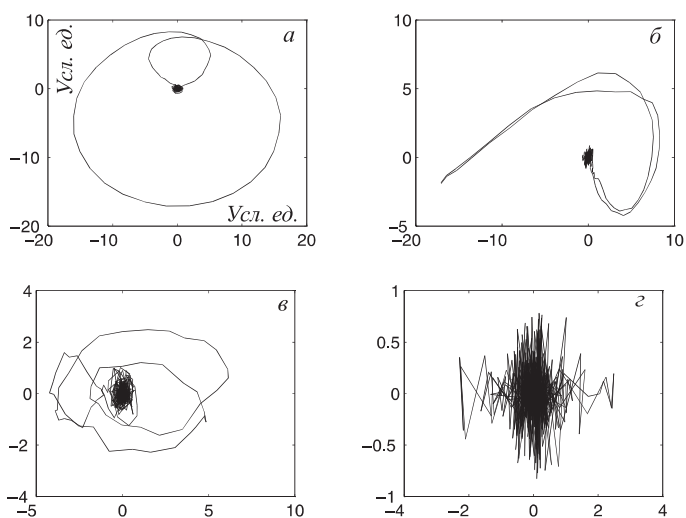


Рис. А.5: Диаграммное представление пар главных компонент: M и $M - 1$ (а); $M - 1$ и $M - 2$ (б); $M - 2$ и $M - 3$ (в); $M - 3$ и $M - 4$ (г). Максимальное число главных компонент — 20.

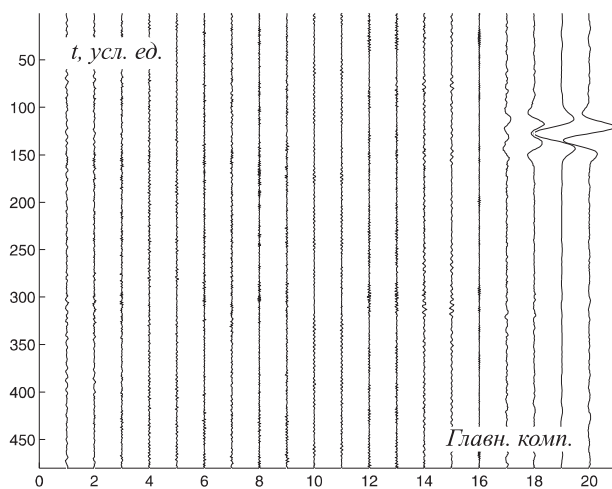


Рис. А.6: Полный набор главных компонент ($M = 20$).

Упражнение

Используя скрипт `р0сер1.m`, сравнить операцию свертки и суммирование в кепстральной области.

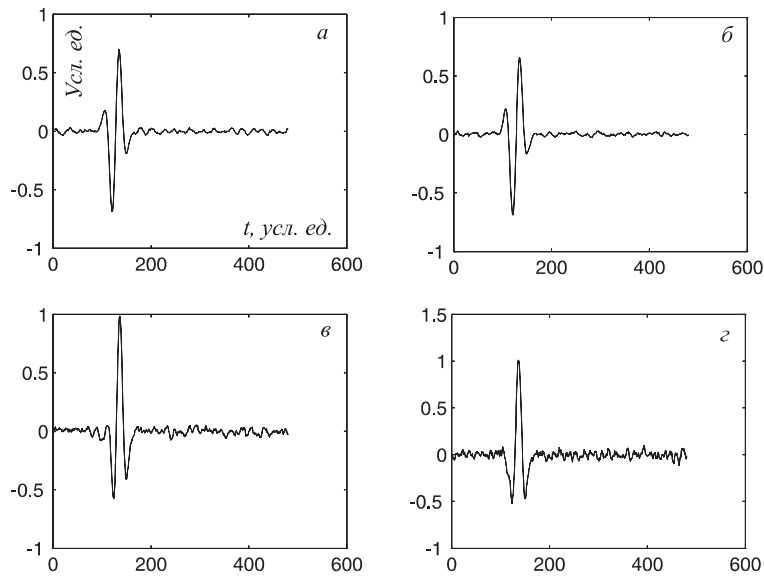


Рис. А.7: Восстановление исходной записи, представленной на рис. А.2, б. Восстановление выполнено только по первой главной компоненте (а), а также суммированием двух (б), трех (в) и четырех (г) главных компонент (см. рис. А.6).

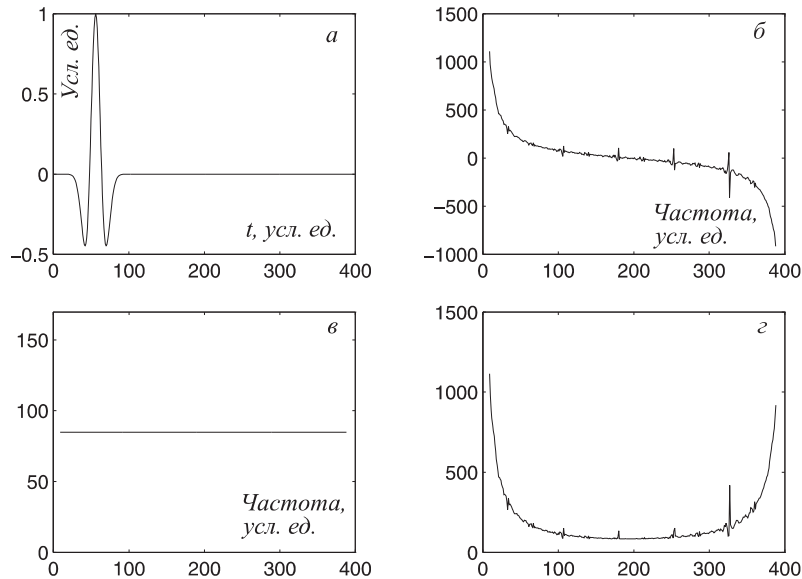


Рис. А.8: Первый сигнал (а) и его кепстральные характеристики (б – г). Импульс Риккера (а), вещественная (б), мнимая (в) части и модуль его кепстра (г).

А.3 Решение прямых и обратных задач

А.3.1 Моделирование гравитационного поля

Гравитационное поле однородной сферы

Скрипт `f0sph1.m` моделирует гравитационное поле однородной сферы.

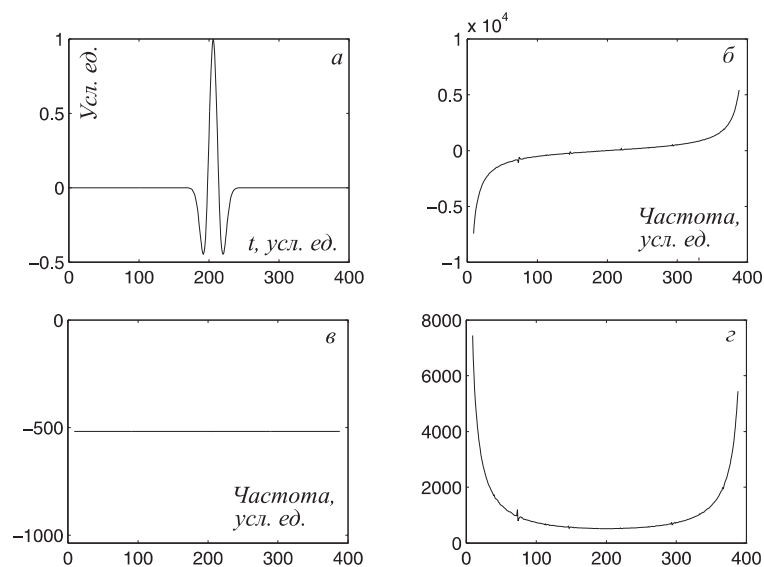


Рис. А.9: Второй сигнал и его кепстральные характеристики (см. рис. А.8).

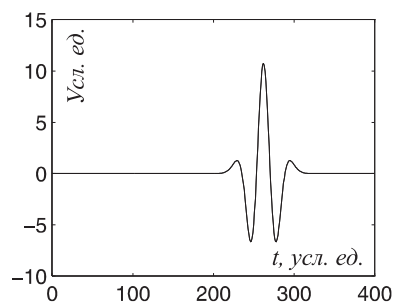


Рис. А.10: Свертка сигналов, представленных на рис. А.8, а, А.9, а и обратное кепстральное преобразование суммы их кепстров.

Гравитационный потенциал V однородной сферы дается формулой

$$V = \gamma \frac{4 \pi a^3 \rho}{3 r}, \quad (\text{A.14})$$

где γ — гравитационная постоянная; a — радиус сферы; ρ — массовая плотность; r — расстояние от центра сферы до точки наблюдения.

Гравитационная сила, порожденная однородной сферой, записывается в виде

$$\vec{g} = -\gamma \frac{4 \pi a^3 \rho}{3 r^2} \vec{n}, \quad (\text{A.15})$$

где \vec{n} — единичный вектор, направленный из центра сферы в точку наблюдения.

Гравитационное поле однородного цилиндра

Скрипт `f0cyl1.m` моделирует гравитационное поле однородного цилиндра.

Гравитационный потенциал V однородного горизонтального цилиндра дается формулой

$$V = 2\pi a^2 \rho \gamma \log \frac{1}{r}, \quad (\text{A.16})$$

где γ — гравитационная постоянная; a — радиус цилиндра; ρ — массовая плотность; r — расстояние между осью цилиндра и точкой наблюдения.

Гравитационная сила, порожденная однородным горизонтальным цилиндром, записывается в виде

$$\vec{g} = -\frac{2\pi a^2 \rho \gamma}{r} \vec{n}, \quad (\text{A.17})$$

где \vec{n} — единичный вектор, направленный из центра цилиндра в точку наблюдения.

А.3.2 Моделирование магнитного поля

Магнитная индукция диполя

Скрипт `f0dip1.m` моделирует магнитную индукцию, вызванную магнитным диполем.

Магнитный потенциал V магнитного диполя дается формулой

$$V = C_m \frac{(\vec{m} \cdot \vec{r})}{r^3}, \quad (\text{A.18})$$

где C_m — равно $\mu_0/4\pi$ в системе СИ; \vec{m} — дипольный момент; r — расстояние между диполем и точкой наблюдения. Магнитная индукция записывается в виде

$$\vec{B} = C_m \frac{m}{r^3} [3(\vec{n}_m \cdot \vec{n}_r) \vec{n}_r - \vec{n}_m], \quad (\text{A.19})$$

где $m = |\vec{m}|$, $\vec{n}_m = \vec{m}/|\vec{m}|$, $\vec{n}_r = \vec{r}/|\vec{r}|$.

Магнитная индукция горизонтального цилиндра

Скрипт `f0mcyl1.m` моделирует магнитную индукцию, вызванную горизонтальным магнитным цилиндром.

Магнитный потенциал горизонтального магнитного цилиндра дается формулой

$$V = 2C_m \frac{(\vec{m} \cdot \vec{n}_r)}{r}, \quad (\text{A.20})$$

где C_m — равно $\mu_0/4\pi$ в системе СИ; \vec{m} — дипольный момент на единицу длины; $r = |\vec{r}|$ — расстояние между центром цилиндра и точкой наблюдения; $\vec{n}_r = \vec{r}/|\vec{r}|$.

Магнитная индукция записывается в виде

$$\vec{B} = \frac{2C_m m}{r^2} [2(\vec{n}_m \cdot \vec{n}_r) \vec{n}_r - \vec{n}_m], \quad (\text{A.21})$$

где $m = |\vec{m}|$, $\vec{n}_m = \vec{m}/|\vec{m}|$.

А.3.3 Моделирование сейсмического поля

Скрипт `fd0m1.m` реализует численное моделирование распространения упругих волн конечно-разностным методом (1-D задача) в полупространстве с внутренними границами (на которых параметры среды испытывают разрыв).

Компонента поля смещений φ удовлетворяет дифференциальному уравнению в частных производных

$$(\lambda + 2\mu)\frac{\partial^2 \varphi}{\partial z^2} + \frac{\partial(\lambda + 2\mu)}{\partial z} \frac{\partial \varphi}{\partial z} + f = \rho \frac{\partial^2 \varphi}{\partial t^2} \quad (\text{A.22})$$

и граничным условиям на свободной границе ($z = 0$)

$$\left. \frac{\partial \varphi}{\partial z} \right|_{z=0} = 0 \quad (\text{A.23})$$

и внутренних границах ($z = z_i$)

$$\varphi|_{z=z_i-0} = \varphi|_{z=z_i+0}, \quad (\lambda + 2\mu) \left. \frac{\partial \varphi}{\partial z} \right|_{z=z_i-0} = (\lambda + 2\mu) \left. \frac{\partial \varphi}{\partial z} \right|_{z=z_i+0}, \quad (\text{A.24})$$

где ρ – массовая плотность, λ и μ – параметры Ламе. Волновое поле возбуждается точечным источником

$$f = \delta(z)\hat{f}(t), \quad (\text{A.25})$$

расположенным в точке $z = z_0$.

Кусочно-однородное полупространство может содержать плавную неоднородность величины $E = \lambda + 2\mu$, которая задается формулой

$$\delta E(z) = \begin{cases} \delta E = 0, & \text{если } |z - \tilde{z}| > \Delta, \\ \delta E = E_m 0.5[1 + \cos(\pi(z - \tilde{z})/\Delta)], & \text{если } |z - \tilde{z}| < \Delta, \end{cases} \quad (\text{A.26})$$

или плавную неоднородность массовой плотности ρ :

$$\delta \rho(z) = \begin{cases} \delta \rho = 0, & \text{если } |z - \tilde{z}| > \Delta, \\ \delta \rho = \rho_m 0.5[1 + \cos(\pi(z - \tilde{z})/\Delta)], & \text{если } |z - \tilde{z}| < \Delta, \end{cases} \quad (\text{A.27})$$

где \tilde{z} – положение центра неоднородности, Δ – полуширина неоднородности, E_m и ρ_m – соответственно максимальные величины возмущения E и ρ .

Упражнения

1. Вычислить волновое поле в кусочно-однородном полупространстве для следующей модели среды и схемы наблюдений:

расположение источника – $z_s = 0.003$ км;

расположение приемников – $z_r = 0.003, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5$ км;

расположение границ – $z=0.3, 0.5, 0.7, 0.9$ км;

упругие параметры — $E = \lambda + 2\mu = 1$;

массовая плотность $\rho = 1.0, 1.2, 0.8, 1.2, 1.0$;

шаг по времени $dt = 0.001$ с;

временная функция источника — импульс Риккера с частотой 25 Гц.

На рис. А.11 представлены модель среды, временная функция в источнике и первая трасса. Выполнить качественный анализ волнового поля. Сейсмограмма

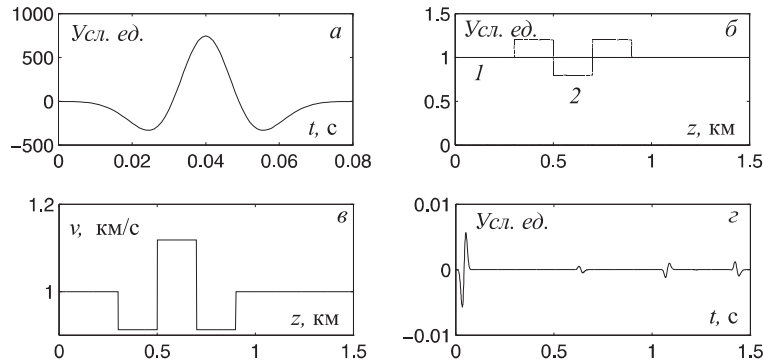


Рис. А.11: Исходные данные для моделирования распространения упругих волн: зависимость воздействия в источнике от времени (а); 1 — ρ , 2 — $\lambda + 2\mu$ (б); скорость распространения продольных волн (в); волновое поле в первой точке наблюдения (г).

представлена на рис. А.12.

А.3.4 Деконволюция с использованием фильтра Винера

Скрипт `rodecon2.m` дает пример деконволюции с использованием винеровского фильтра ³.

Дискретная свертка векторов \vec{h} и \vec{x} (см. 3.4) определяется следующим образом

$$y_k = \sum_{j=0}^{L_h} h_j x_{k-j}, \quad k = 0, \dots, L_y \quad (L_y = L_h + L_x - 1), \quad (\text{A.28})$$

где L_h — длина фильтра \vec{h} , L_x — длина входного сигнала (вектора на входе) \vec{x} . Дискретная свертка может быть представлена в матричной форме

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{L_y} \end{bmatrix} = \begin{bmatrix} x_0 & 0 & 0 & \dots & 0 \\ x_1 & x_0 & 0 & \dots & 0 \\ x_2 & x_1 & x_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & x_{L_x} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{L_h} \end{bmatrix} \quad (\text{A.29})$$

³В свою очередь фильтр Винера можно рассматривать как пример применения метода наименьших квадратов (см. 6.1).

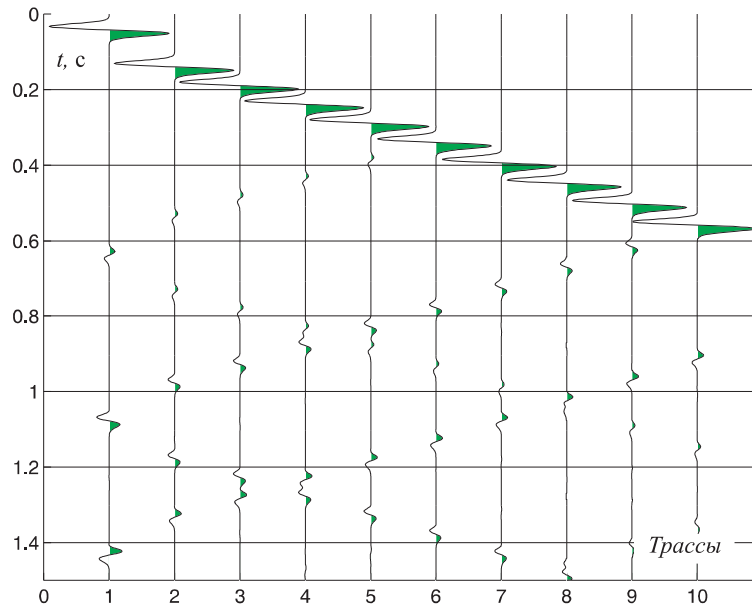


Рис. А.12: Сейсмограмма.

или

$$\vec{y} = X\vec{h}. \quad (\text{A.30})$$

Если \vec{y} и X – известные величины, используя метод наименьших квадратов, неизвестная величина \vec{h} может быть найдена решением системы линейных уравнений

$$X^T X \vec{h} = X^T \vec{y} \quad (\text{A.31})$$

или, вводя регуляризующий член,

$$(X^T X + \alpha I) \vec{h} = X^T \vec{y}, \quad (\text{A.32})$$

где I – единичная матрица, α – регуляризующий коэффициент.

Если только один элемент вектора \vec{y} равен 1, а остальные элементы равны 0, то filter \vec{h} is named the Wiener filter. In this case \vec{x} is transformed to $\vec{y} = [\dots, 1, \dots]$ by the filter \vec{h} .

Рассмотрим пример работы скрипта `r0decon2.m`. Пусть имеется сейсмограмма из п. А.3.3 (см. упражнение к п. А.3.3, рис. А.12) с дополнительным шумом. Поставим задачу уменьшить отношение шум/сигнал и приблизить форму сигналов на сейсмограмме к δ -функции. На рис. А.13 представлен фрагмент сейсмограммы из рис. А.12 с добавлением шума с нормальным распределением ($N(0, \sigma A_m)$, где $\sigma=0.1$, A_m – максимальная величина сейсмического сигнала на сейсмограмме). На рис. А.14 изображен фрагмент сейсмограммы, представленной на рис. А.13, после деконволюции.

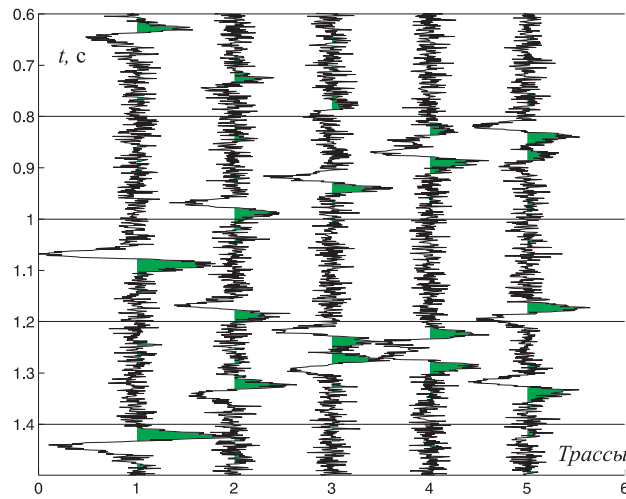


Рис. А.13: Фрагмент сейсмограммы из рис. А.12 с добавлением некоррелированного шума с нормальным распределением.

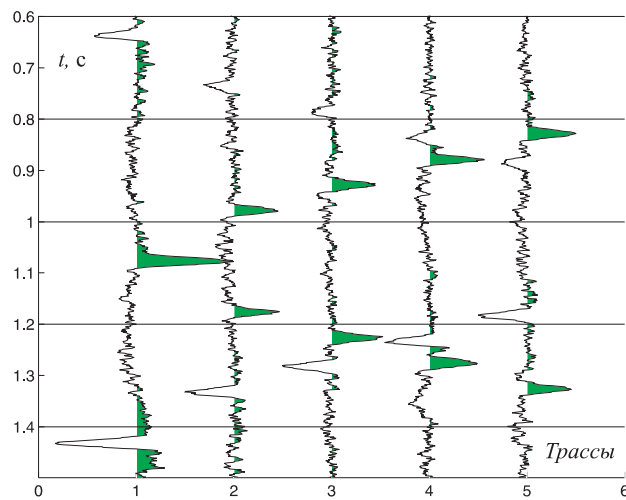


Рис. А.14: Деконволюция сейсмограммы, представленной на рис. А.13.

Упражнения

Выполнить деконволюцию сейсмического поля:

без добавления шума (`sig1=sig2=0`);

при различных значениях регуляризующего коэффициента α (см. формулу (A.32))
(при `eps1=2.0, 1.0, 0.5`);

при различных величинах уровня шума (`sig1=sig2=0.5, 0.1`).

А.3.5 Количественная интерпретация

Рассмотрим применение метода Ньютона–Лекама (см. 5.2, 6.5, 7.1) для точечного оценивания параметров $\vec{\theta}$ геофизических объектов (script `p0mg1.m`). В качестве примера рассмотрим тела с вертикальной намагниченностью:

нить полюсов

$$f_k(\vec{\theta}) = \frac{2Mh}{h^2 + (x_k - \xi)^2} \quad \vec{\theta} = \{M, h, \xi\}; \quad (\text{A.33})$$

нить диполей

$$f_k(\vec{\theta}) = \frac{2M[h^2 - (x_k - \xi)^2]}{[h^2 + (x_k - \xi)^2]^2}, \quad \vec{\theta} = \{M, h, \xi\}; \quad (\text{A.34})$$

полюс

$$f_k(\vec{\theta}) = \frac{M(2h^2 - x_k^2)}{(h^2 + x_k^2)^{5/2}}, \quad \vec{\theta} = \{M, h\}; \quad (\text{A.35})$$

диполь

$$f_k(\vec{\theta}) = \frac{Mh}{(h^2 + x_k^2)^{3/2}}, \quad \vec{\theta} = \{M, h\}; \quad (\text{A.36})$$

где M — магнитный момент, h — глубина, ξ — положение максимума функции $f_k(\vec{\theta})$.

Применяя метод максимального правдоподобия к аддитивной модели экспериментального материала

$$u_k = f_k(\vec{\theta}) + \varepsilon_k \quad (\vec{u} = \vec{f}(\vec{\theta}) + \vec{\varepsilon}), \quad (\text{A.37})$$

содержащей некоррелированную нормальную шумовую компоненту ε_k ($N(0, \sigma)$), можно получить следующую процедуру оценивания вектора искомых величин $\vec{\theta} = (\theta_1, \dots, \theta_S)$. Сначала выбирается вектор начальных значений $\vec{\theta}_0$ для вектора искомых величин $\vec{\theta}$, затем решается система линейных уравнений (см. 7.1)

$$\tilde{C} \Delta \vec{\theta}^{(1)} = \vec{d}, \quad (\text{A.38})$$

где

$$d_s = \frac{1}{\sigma^2} \left[(\vec{u} - \vec{f}(\vec{\theta}))^T \frac{\partial \vec{f}(\vec{\theta})}{\partial \theta_s} \right] \Bigg|_{\vec{\theta} = \vec{\theta}^{(0)}},$$

$$\tilde{c}_{ss'} = \frac{1}{\sigma^2} \frac{\partial f^T(\vec{\theta})}{\partial \theta_s} \Bigg|_{\vec{\theta} = \vec{\theta}^{(0)}} \frac{\partial f(\vec{\theta})}{\partial \theta_{s'}} \Bigg|_{\vec{\theta} = \vec{\theta}^{(0)}},$$

$$s = 1, \dots, S$$

и вычисляется первое приближение $\hat{\theta}^{(1)}$ для искомых величин $\vec{\theta}$

$$\hat{\theta}^{(1)} = \vec{\theta}^{(0)} + \Delta\vec{\theta}^{(1)}. \quad (\text{A.39})$$

Итерационный процесс заканчивается, если выполнены пороговые условия, например,

$$|\Delta\hat{\theta}_s^{(n)} / \hat{\theta}_s^{(n)}| \approx 10^{-2} \div 10^{-3}. \quad (\text{A.40})$$

Пример подобного численного моделирования для поля из формулы (А.33) представлен на рис. А.15 для следующего набора входных данных:

`x=-4:0.1:4;` % точки наблюдения

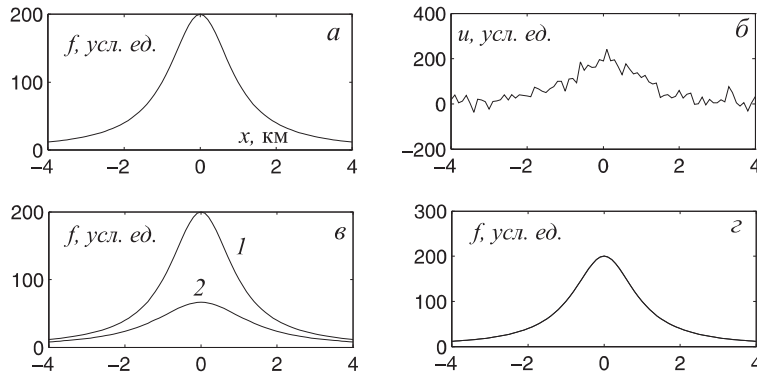


Рис. А.15: Иллюстрация количественной интерпретации. Модельное поле без шума (а), с шумом (б), модельное поле (1) и поле (2), отвечающее начальным параметрам (в), сравнение модельного поля и поля объекта с параметрами, полученными в результате оценивания (г).

```

m=1; % магнитный момент
h=1; % глубина
xi=0.0; % расположение восстанавливаемого объекта
xm=0.0; % математическое ожидание шума
s=0.1; % стандартное отклонение шума
m0=0.5; % начальная величина m
h0=1.5; % начальная величина h
xi=0.; % начальная величина xi
eps_m=1.0e-5; % пороговое условие для m
eps_h=1.0e-5; % пороговое условие для h
    
```

Выходными данными скрипта `p0mg1.m` являются оценки величин M и h , матрица Фишера и обратная к ней матрица.

Упражнения

Используя скрипт `p0mg1.m`, выполнить следующие задания для моделей (А.33) – (А.36) (обратить внимание на величины элементов матрицы Фишера и обратной

к ней матрицы):

1. Найти оценки параметров M и h при различных величинах стандартного отклонения (например, $\sigma = 0.1, 0.2, 0.3$).

2. Исследовать сходимость итерационного процесса при различных величинах начальных значений M и h .

3. Выполнить моделирование при различном числе точек наблюдения (например 10, 20, 50, 100).

А.3.6 Качественная интерпретация

Скрипт `r0mg2.m` реализует алгоритм качественной интерпретации (см. 5.2) — “сигнал и шум (H_0) или шум (H_1)”, основанный на критерии отношения апостериорных вероятностей (см. 8, 8.2) для моделей сигналов из формул (А.33)–(А.36). В случае некоррелированного нормального шума такой критерий записывается в следующем виде:

$$\alpha = \ln \frac{P(1)}{P(0)} + \frac{1}{2\sigma^2} \left[\vec{u}^T \vec{u} - (\vec{u} - \vec{f}_1)^T (\vec{u} - \vec{f}_1) \right]. \quad (\text{А.41})$$

Если $\alpha \geq 0$, то принимается гипотеза H_1 (“шум”). Если $\alpha < 0$, то принимается гипотеза H_0 (“сигнал и шум”).

Предполагается, что модельная функция $f_1(\vec{\theta})$ отличается от “истинной” функции $f_1(\vec{\theta} + \Delta\vec{\theta})$. Такое отличие задается отклонением $\Delta\vec{\theta}$ “истинных” параметров $\vec{\theta} + \Delta\vec{\theta}$ от их предполагаемых значений $\vec{\theta}$.

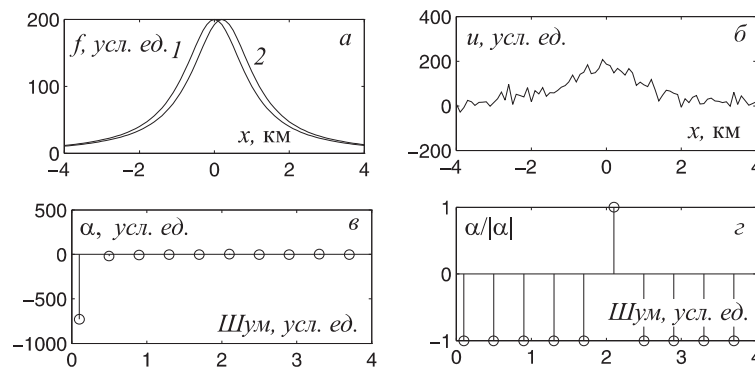


Рис. А.16: Иллюстрация качественной интерпретации. Модельное поле $f_1(\vec{\theta})$ (1) и “истинное” поле $f_1(\vec{\theta} + \Delta\vec{\theta})$ (2) (а), модельное поле с некоррелированным шумом (б), зависимость параметра α (в) и нормированного параметра $\alpha/|\alpha|$ (г) от величины шума.

Пример качественной интерпретации для поля (А.33) демонстрируется на рис. А.16 при следующих значениях входных данных:

```

x=-4:0.1:4; % точки наблюдения
m=1; % магнитный момент
h=1; % глубина
xi=0.0; % положение объекта
ddm=0.0; % отклонение m
ddh=0.0; % отклонение h
ddxi=0.2; % отклонение xi
xm=0.0; % математическое ожидание шума
p0=0.5; % априорная вероятность существования объекта
s=0.1; % начальное стандартное отклонение для шума
s1=s:0.4:4; % величины стандартных отклонений для шума

```

Упражнения

Используя скрипт `p0mg2.m`, выполнить для моделей (А.33) – (А.36):

- 1) моделирование для широкого диапазона изменения отношения шум/сигнал (переменная `s1`);
- 2) моделирование для различных величин расхождения $\Delta\vec{\theta}$ аргументов модельной функции и “истинного” поля;
- 3) моделирование для различных величин априорной вероятности, например, $p_0=0,25; 0,5; 0,75$;
- 4) численное моделирование для различного числа точек наблюдения (например 10, 20, 50, 100).

А.3.7 Дифракционная томография

Скрипт `dt0mn1.m` реализует численное моделирование восстановления (определения параметров) неоднородности среды методом дифракционной томографии (см. 10).

Прямая задача на распространение упругих волн (1-D) в однородном пространстве, содержащем плавное возмущение параметров с размером порядка длины волны, решается конечно-разностным методом. При этом компонента волнового поля φ удовлетворяет дифференциальному уравнению в частных производных (А.22) и граничным условиям (А.23), (А.24). Восстановление возмущений параметров λ , μ и ρ осуществляется методом дифракционной томографии по следующему алгоритму.

Если ввести томографические функционалы p_ρ и p_E , связанные с величинами ρ и $E = \lambda + 2\mu$ соответственно:

$$p_\rho(z, z_s, z_r, t) = - \int_0^\infty \varphi_{\text{out}}(z, z_r, t - \tau) \frac{\partial^2}{\partial t^2} \varphi_{\text{in}}(z, z_s, \tau) d\tau,$$

$$\begin{aligned}
p_E(z, z_s, z_r, t) &= - \int_0^\infty \frac{\partial}{\partial z} \varphi_{\text{out}}(z, z_r, t - \tau) \times \\
&\times \frac{\partial}{\partial z} \varphi_{\text{in}}(z, z_s, \tau) d\tau,
\end{aligned} \tag{A.42}$$

и использовать приближение Борна, то рассеянное (или *разностное*) поле $\Delta\varphi$, порожденное неоднородностью, можно записать в виде

$$\Delta\varphi \approx \int_L (p_E \delta E + p_\rho \delta \rho) dl, \tag{A.43}$$

где z_s и z_r — положение источника и точки наблюдения; φ_{in} — волновое поле, генерируемое источником; φ_{out} — волновое поле, генерируемое расположенным в точке наблюдения “искусственным” источником с временной зависимостью $\delta(t)$; L — область восстановления параметров среды; δE , $\delta \rho$ — искомые величины (возмущение E и ρ относительно опорной среды).

Если, например, считается выполненным $\delta \rho \approx 0$ и определению подлежит только δE , то интегральное уравнение (A.43) после дискретизации может быть представлено как система линейных уравнений

$$P \vec{\delta}_E = \vec{\delta}_\varphi, \tag{A.44}$$

где $\vec{\delta}_E$ — искомый вектор, а $\vec{\delta}_\varphi$ — разностное поле.

Одна из простых схем регуляризации решения системы линейных уравнений (A.44) состоит в решении следующей системы линейных уравнений:

$$(P'P + \varepsilon D'D) \vec{\delta}_E = P' \vec{\delta}_\varphi, \tag{A.45}$$

где ε — параметр регуляризации, D — например, единичная матрица.

Рассмотрим пример численного моделирования для следующей модели:

$$z_s = z_r = 0.003 \text{ км};$$

$$\rho = 1, E = 1, v = \sqrt{E/\rho} = 1;$$

искомая неоднородность задается формулой

$$\begin{aligned}
\delta E(z) &= \\
&= \begin{cases} \delta E = 0, & \text{если } |z - \tilde{z}| > \Delta, \\ \delta E = E_m 0.5 [1 + \cos(\pi(z - \tilde{z})/\Delta)], & \text{если } |z - \tilde{z}| < \Delta, \end{cases}
\end{aligned} \tag{A.46}$$

где \tilde{z} — расположение центра неоднородности, Δ — полуширина неоднородности, E_m — максимальная величина возмущения E .

$\tilde{z} = 0.3$ км, $\Delta = 0.03$ км, $E_m = 0.3$; импульс Риккера

$$F(\omega) = \left(\frac{\omega}{\omega_0}\right)^2 \exp\left\{\left(\frac{\omega}{\omega_0}\right)^2\right\} \exp\{-2\pi i \omega / \omega_0\} \tag{A.47}$$

с доминирующей частотой $\omega_0 = 25$ Гц; интервал L восстановления параметров среды $L = 0.21 - 0.39$ км; параметр регуляризации $\varepsilon = 10^{-5}$.

Графическое представление модели среды и результата решения прямой задачи дано на рис. А.17 и рис. А.18 соответственно. Результат восстановления

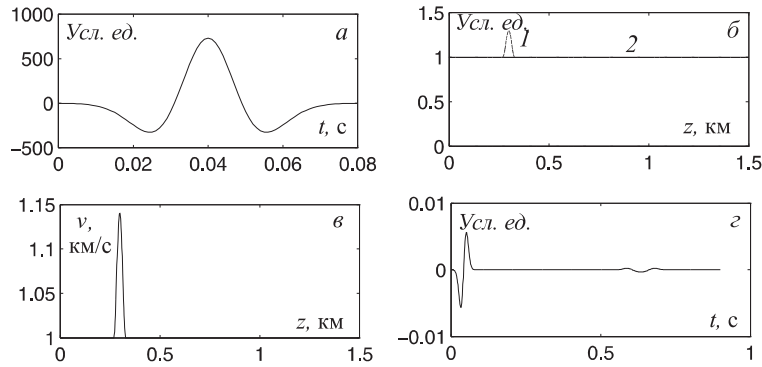


Рис. А.17: Исходные данные для моделирования распространения упругих волн и волновое поле в первой точке наблюдения. Форма представления данных та же, что и на рис. А.11.

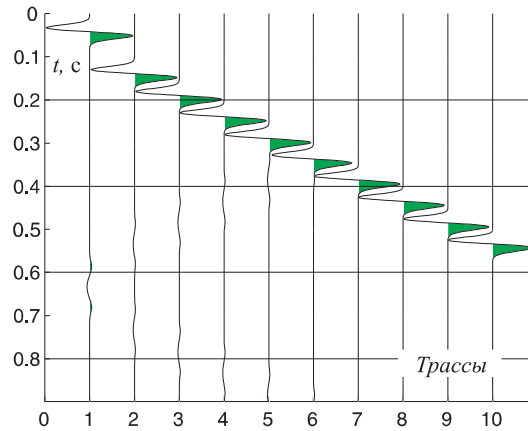


Рис. А.18: Сейсмограмма для точек наблюдения (км): 0.003, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5.

параметров неоднородности представлен на рис. А.19.

Упражнения

1. Используя скрипт `dt0mn1`, исследовать точность восстановления параметра E методом дифракционной томографии в зависимости от контрастности и размера неоднородности (например, $E = 0.05, 0.1, 0.2, 0.5$ $\Delta = 0.01, 0.02, 0.03, 0.04$).

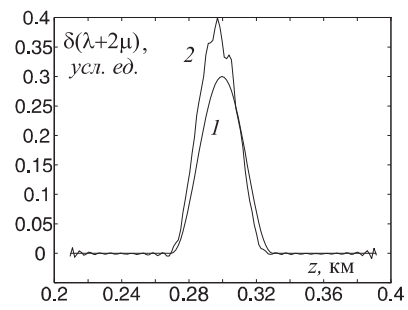


Рис. А.19: Восстановление возмущения скорости распространения продольной волны: 1 — модель неоднородности; 2 — результат восстановления.

2. Оценить влияние шумовых отклонений в исходных данных на качество восстановления неоднородности.

Оглавление

1	Введение в теорию вероятностей	1
1.1	Определение вероятности	1
1.1.1	Множество элементарных событий	1
1.1.2	Аксиоматическое определение вероятности.	5
1.1.3	Основные свойства вероятности	6
1.1.4	Частотное определение вероятности	6
1.1.5	Классическое определение	7
1.1.6	Геометрическое распределение	7
1.1.7	Условная вероятность и независимость случайных событий	7
1.1.8	Формула полной вероятности	8
1.1.9	Теорема Байеса	9
1.2	Функция распределения	9
1.2.1	Случайная переменная	9
1.2.2	Определение функции распределения	10
1.2.3	Свойства функции распределения	10
1.2.4	Плотность вероятностей	11
1.2.5	Функция распределения и плотность вероятностей для функ- ции, зависящей от случайной переменной	12
1.2.6	Случайные векторы	13
1.2.7	Маргинальные и условные распределения	14
1.2.8	Совместные распределения и плотности вероятностей	15
1.3	Математическое ожидание и моменты случайных величин	17
1.3.1	Математическое ожидание	17
1.3.2	Дисперсия	19
1.3.3	Свойства дисперсии и коэффициентов корреляции	19
1.3.4	Квантили	21
1.3.5	Характеристики плотности вероятности	21
1.4	Характеристические и производящие функции	24
1.4.1	Характеристические функции	24
1.4.2	Производящие функции моментов	25
1.4.3	Семиинварианты или кумулянты	26
1.5	Предельные теоремы	26
1.5.1	Сходимость по вероятности	26
1.5.2	Неравенство Чебышева	27

1.5.3	Теорема Чебышева	27
1.5.4	Центральная предельная теорема	28
1.6	Функции распределения дискретных случайных величин	29
1.6.1	Биномиальное распределение	30
1.6.2	Распределение Пуассона	31
1.6.3	Геометрическое распределение	32
1.7	Непрерывные случайные величины	33
1.7.1	Одномерное нормальное распределение	33
1.7.2	Многомерное нормальное распределение	35
1.7.3	Равномерное распределение	37
1.7.4	Распределение χ^2	38
1.7.5	Распределение Стьюдента (t-распределение)	39
1.7.6	Распределение Фишера	40
1.7.7	Экспоненциальное распределение	41
1.7.8	Распределение Лапласа	41
1.7.9	Распределение Коши	42
1.7.10	Логарифмически нормальное распределение	43
1.7.11	Значение нормального распределения	43
1.7.12	Доверительные интервалы	44
2	Информация и энтропия	47
2.1	Энтропия множества дискретных состояний системы	47
2.2	Энтропия сложной системы	47
2.3	Информация Шеннона (дискретный случай)	49
2.4	Непрерывный случай (энтропия и информация)	50
2.5	Информация Фишера	54
3	Свойства случайных функций	57
3.1	Математическое ожидание и дисперсия	57
3.2	Свойства случайных функций	58
3.2.1	Стационарность	58
3.2.2	Эргодичность	59
3.3	Свойства автокорреляционных функций	59
3.4	Действие линейного оператора на случайную функцию	60
3.5	Взаимокорреляционная функция	61
3.6	Теорема Винера–Хинчина и спектр мощности	62
3.7	Пример корреляционных функций и спектров мощности	63
3.8	Оценивание численных характеристик случайных функций	65
4	Элементы математической статистики	67
4.1	Оценивание	67
4.2	Состоятельные оценки	68
4.3	Несмещенное оценивание	69
4.4	Неравенство Рао–Крамера. Эффективность	70
4.5	Достаточность	72

4.6	Робастное оценивание	72
5	Модели измерений	73
5.1	Аддитивные модели	73
5.2	Модели количественной интерпретации	74
5.3	Регрессионная модель	75
5.4	Модели количественной интерпретации	77
5.5	Модели качественно-количественной интерпретации	78
5.6	Случайная компонента модели и ее свойства	79
5.6.1	Нормальное распределение	79
5.6.2	Случайная компонента как стационарный процесс	79
5.6.3	Случайная компонента как некоррелированный процесс	80
5.6.4	Случайная компонента как марковский процесс	80
5.6.5	Распределение Лапласа	81
5.7	Модель со случайным параметром	81
5.8	Априорная информация о поле параметров	82
6	Обратные задачи: линейные модели	85
6.1	Метод наименьших квадратов (МНК)	85
6.1.1	Свойства МНК оценок	86
6.2	Метод наименьших квадратов: нелинейная модель	87
6.3	Метод наименьших квадратов: ортогональные полиномы	88
6.4	Метод наименьших квадратов с ограничениями	90
6.5	Метод максимального правдоподобия (ММП)	93
6.6	Метод наименьших модулей	95
6.7	Робастный метод Хьюбера	96
6.8	Робастный метод Эндрюса	97
6.9	Метод Байеса и статистическая регуляризация	99
6.9.1	Рекурсивный алгоритм для метода максимума апостериорной вероятности	101
6.9.2	Фильтр Калмана	103
6.10	Сингулярный анализ и метод наименьших квадратов	104
6.10.1	Матрица разрешения	108
6.11	Метод Бейкуса–Гильберта	110
7	Обратные задачи: нелинейные модели	113
7.1	Метод Ньютона Лекама	113
7.2	Метод наименьших квадратов: нелинейный случай	115
7.3	Метод наименьших квадратов для нестационарной модели	116
7.4	Метод интервального оценивания	118
7.5	Генетический алгоритм	119
7.5.1	Кодирование	120
7.5.2	Отбор	120
7.5.3	Скращивание	121
7.5.4	Мутация	122

7.5.5	Выбор	122
8	Статистические критерии выбора модели	127
8.1	Параметрические критерии выбора модели	127
8.2	Критерий отношения апостериорных вероятностей	128
8.3	Задача разделения сигналов	131
9	Геофизическая лучевая томография	145
9.1	Основные определения	145
9.2	Преобразование Радона	145
9.3	Алгебраические методы	147
9.4	Восстановление медленности и коэффициента поглощения	150
9.5	Преобразование Радона сейсмограмм	152
9.6	Преобразование Гильберта и аналитический сигнал	154
10	Дифракционная томография	157
10.1	Дифракционная томография в случае скалярного волнового уравнения	157
10.2	Модель сейсмической трассы в приближении Борна	161
10.3	Дифракционная томография в случае векторного поля смещений	163
10.3.1	Модель измерений	163
10.3.2	Томографический функционал	165
10.3.3	Томографический функционал для скалярного волнового уравнения	167
10.3.4	Томографический функционал для уравнения Ламе	167
10.4	Алгоритм реконструкции	173
10.5	Реконструктивная томография и метод Бейкуса–Гильберта	176
11	Многофакторный анализ временных рядов	181
A	Компьютерные упражнения	187
A.1	Статистические методы	187
A.1.1	Численное моделирование случайных величин	187
A.1.2	Построение гистограмм	188
A.1.3	Описание случайных переменных	189
A.1.4	Генераторы случайных чисел	195
A.1.5	Доверительные интервалы	196
A.1.6	Временные ряды	197
A.2	Преобразования	197
A.2.1	Преобразование Фурье	197
A.2.2	Простейшие сигналы и их спектральные характеристики	197
A.2.3	Многофакторный анализ	198
A.2.4	Кепстральное преобразование	200
A.3	Решение прямых и обратных задач	203

А.3.1	Моделирование гравитационного поля	203
А.3.2	Моделирование магнитного поля	205
А.3.3	Моделирование сейсмического поля	206
А.3.4	Деконволюция с использованием фильтра Винера	207
А.3.5	Количественная интерпретация	210
А.3.6	Качественная интерпретация	212
А.3.7	Дифракционная томография	213