МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ В.Ф. УТКИНА

Кафедра радиоуправления и связи

МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

по дисциплине (модулю)

«Цифровые многоканальные системы передачи информации»

Направление подготовки

11.05.01 «Радиоэлектронные системы и комплексы»

Направленность (профиль) подготовки

Радиосистемы и комплексы управления

Уровень подготовки специалитет

Программа подготовки специалитет

Квалификация выпускника – инженер Форма обучения – очная

Рязань 2025

5244

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

ЦИФРОВАЯ ОБРАБОТКА СЛУЧАЙНЫХ СИГНАЛОВ

Методические указания к курсовой работе

Рязань 2018

УДК 621.396.21
Цифровая обработка случайных сигналов: методические указания к курсовой работе / Рязан. гос. радиотехн. ун-т; сост. В.В. Езерский. Рязань, 2018. 24 с.

Содержат краткие теоретические сведения по теории оценок параметров и характеристик случайных сигналов, задание на выполнение и требования по оформлению.

Предназначены для бакалавров ООП «Многоканальные телекоммуникационные системы» направления 11.03.02 «Инфокоммуникационные технологии и системы связи» и магистрантов ООП «Многоканальные телекоммуникационные системы» направления 11.04.02 «Инфокоммуникационные технологии и системы связи» факультета радиотехники и телекоммуникаций.

Табл. 2. Ил. 6. Библиогр.: 4 назв.

Метод максимального правдоподобия, оценка моментов, законов распределения плотности вероятности, корреляционной функции и спектральной плотности мощности

Печатается по решению редакционно-издательского совета Рязанского государственного радиотехнического университета.

Рецензент: кафедра радиоуправления и связи Рязанского государственного радиотехнического университета (зав. кафедрой С.Н. Кириллов)

Цифровая обработка случайных сигналов

Составитель Е з е р с к и й Виктор Витольдович

Редактор Р.К. Мангутова
Корректор С. В. Макушина
Подписано в печать 24.04 18. Формат бумаги 60x84 1/16.
Бумага писчая. Печать трафаретная. Усл. печ. л. 1,5.
Тираж 150 экз. Заказ
Рязанский государственный радиотехнический университет.
390005, Рязань, ул. Гагарина, 59/1.
Редакционно-издательский центр РГРТУ.

Введение

С обработкой случайных сигналов приходится иметь дело в любой радиотехнической системе. Радиотехнические системы являются информационными, сигнал в них переносит информацию. Поэтому важнейшая задача – это извлечение информации из сигнала. С этой целью необходимо оценивать параметр сигнала, в котором содержится информация. Наиболее типичной задачей обработки является оценка различных характеристик случайного сигнала (математическое ожидание, дисперсия, корреляционная функция, законы распределения, спектральная плотность мощности и т.д.). Исходными данными для оценки является массив выборочных значений $x_1, x_2, \ldots, x_N$. Здесь N – размер выборки.

Далее ограничимся рассмотрением стационарного эргодического процесса.

Существуют различные методы нахождения оценок. Одной из наиболее полезных оценок является оценка максимального правдоподобия.

Оценка максимального правдоподобия и её характеристики

Оценка максимального правдоподобия параметра сигнала a_x основана на анализе совместного закона плотности вероятности выборки, который называют функцией правдоподобия:

$$W_N(x_1, x_2, \ldots, x_N \mid a_x),$$

где $W_N(\bullet)$ – N-мерный совместный закон плотности вероятности выборки; a_x – параметр, который требуется оценить.

Оценка максимального правдоподобия – это такое значение a_x, при котором функция правдоподобия $W_N(x_1, x_2, \ldots, x_n \mid a_x)$ становится максимальной.

Уравнение для нахождения оценки максимального правдоподобия имеет вид:

$$\frac{dW_N(x_1, x_2, \ldots, x_N \mid a_x)}{da_x} = 0.$$

Это уравнение называется уравнением максимального правдоподобия. Решением этого уравнения является оценка максимального правдоподобия $a_x = F(x_1, x_2, \ldots, x_N \mid a_x)$, которая является некоторой сложной функцией, зависящей от вида закона распределения.

Вычисление этой оценки существенно упрощается, если элементы выборки $x_1, x_2, \ldots, x_N$ - независимые случайные величины. Для этого специально организуют экспериментальные измерения так, чтобы обеспечить такую независимость. Тогда многомерный закон распределения можно записать в виде произведения одномерных законов:

$$W_N(x_1, x_2, \ldots, x_N \mid a_x) = W_1(x_1 \mid a_x) W_1(x_2 \mid a_x) \ldots = \prod_{i=1}^{N} W_1(x_i \mid a_x) .$$

При этом в уравнении правдоподобия надо вычислять производную от произведения одномерных законов. Это сложно математически и приводит к громоздким выражениям, из которых практически невозможно получить решение. Для облегчения поиска максимально правдоподобной оценки часто ищут максимум не функции правдоподобия, а её логарифма. При этом функция правдоподобия имеет вид:

$$\Lambda(x_1, x_2, \ldots, x_N \mid a_x) = \ln W_N(x_1, x_2, \ldots, x_N \mid a_x) .$$

В этом случае уже приходится вычислять производную от суммы функций, что, несомненно, проще:

$$\frac{d\Lambda(x_1, x_2, \ldots, x_N \mid a_x)}{da_x} = 0 .$$

Так как выборка содержит случайные значения, то любая оценка является случайной величиной. Важными характеристиками любой оценки являются математическое ожидание и дисперсия.

Если математическое ожидание оценки совпадает с точным значением этой величины

$$\mathrm{E}[\hat{a}_x] = a_x ,$$

где $\mathrm{E}[\bullet]$ означает усреднение, то оценка называется несмещённой. В противном случае оценка называется смещённой.

Количественной мерой разброса случайной оценки относительно точного значения является среднеквадратическое отклонение оценки.

Квадрат этой величины называют дисперсией:

$$\mathrm{Var}[\hat{a}_x] = E\left[\left\{\hat{a}_x - E[\hat{a}_x]\right\}^2\right].$$

Физически дисперсия – это мощность переменной составляющей случайного процесса.

Другой мерой качества оценки, связанной с дисперсией оценки, является закон распределения плотности вероятности оценки. Например, для оценки математического ожидания возможны варианты плотности вероятности, показанные на рис. 1.

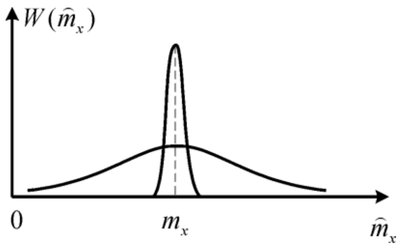

Рисунок 1 – Возможные формы закона распределения оценки

Если функция плотности вероятности широкая, то возможные значения оценки сильно разбросаны вокруг точного значения оцениваемого параметра. Принято считать, что это плохая оценка. И наоборот, если закон распределения сконцентрирован вокруг точного значения, то это оценка хорошая. Чем меньше дисперсия оценки, тем уже закон распределения.

Один из способов описания качества оценки – это указание доверительного интервала, т.е. интервала значений вблизи точного значения, в который оценка попадает с заданной вероятностью. Эту вероятность называют доверительной. Чем этот интервал уже, тем лучше оценка.

В зависимости от того, как ведут себя закон распределения и дисперсия оценки при увеличении объема выборки, различают состоятельную и несостоятельную оценки. Если при росте объема выборки дисперсия оценки стремится к нулю, т.е. закон распределения сужается, стягиваясь к дельта-функции, то оценка состоятельная. Если же этого не наблюдается, то это несостоятельная оценка.

Оценка моментов

Основными моментами, представляющими практический интерес, являются математическое ожидание и дисперсия.

Оценка математического ожидания

По определению математическим ожиданием стационарного процесса x является величина

$$m_x = E[x] = \int_{-\infty}^{\infty} xW(x)dx,$$

где $W(x)$ – функция плотности распределения вероятности.

Это первый начальный момент. Для стационарного эргодического процесса может быть записано

$$m_x = \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{T} x(t)dt.$$

Для дискретного стационарного эргодического сигнала можно записать

$$m_x = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x_n.$$

Если этот метод максимального правдоподобия применить к нормальному закону распределения, то получаем максимально правдоподобную оценку математического ожидания в виде

$$\hat{m}_x = \frac{1}{N} \sum_{n=1}^{N} xn.$$

Эту величину называют выборочным средним.

N здесь должно быть достаточно велико.

При этом $\hat{m}_x$ является случайной величиной, которая имеет свой закон распределения и свои моменты.

Для выборочного среднего оценка является несмещённой:

$$E\left[\frac{1}{N}\sum_{n=1}^{N} x_n\right] = \frac{1}{N}\sum_{n=1}^{N} E[x_n] = \frac{1}{N}\sum_{n=1}^{N} m_x = \frac{1}{N} m_x N = m_x,$$

т.е. указанная оценка математического ожидания является несмещённой.

Дисперсия оценки в виде выборочного среднего равна

$$\mathrm{Var}[\hat{m}_x] = \frac{\sigma_x^2}{N}.$$

При увеличении объёма выборки дисперсия оценки математического ожидания стремится к нулю. Т.е. это состоятельная оценка.

Оценка дисперсии

Определением дисперсии является выражение

$$D_x = \sigma_x^2 = \mathrm{var}[x] = \int_{-\infty}^{\infty} (x - m_x)^2 W(x)dx.$$

Это второй центральный момент.

Для оценки дисперсии нормального стационарного эргодического случайного процесса используется выражение

$$D_x = \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{T} \left[x(t) - m_x\right]^2 dt .$$

Для дискретного стационарного эргодического сигнала можно записать

$$D_x = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \left(x_n - m_x\right)^2 .$$

Отсюда можно записать выражение для выборочной дисперсии:

$$\hat{D}x = \hat{\sigma}_x^2 = \frac{1}{N} \sum (xn - \hat{m}_x)2 \ .$$

Математическое ожидание для этой оценки имеет вид

$$E[\hat{\sigma}_x^2] = \frac{N-1}{N} \sigma_x^2 .$$

Следовательно, оценка дисперсии является смещенной, но при росте N это смещение стремится к нулю, поэтому эту оценку называют асимптотически не смещенной.

Так как нам известен коэффициент искажений, то оценку можно скорректировать:

$$\hat{\sigma}_x^2 = \frac{1}{N-1} \sum (x_n - \hat{m}_x)^2 \ .$$

Это уже несмещенная оценка.

Дисперсия оценки дисперсии имеет вид

$$\mathrm{Var}[\sigma_x^2] = \frac{1}{\mathrm{N}} [E[x^4] - \frac{1}{N} (E[x^2])^2] .$$

Отсюда следует, что она состоятельная, так как при увеличении объёма выборки N дисперсия стремится к нулю.

Учёт особенностей алгоритмов численной оценки

Алгоритмы вычисления математического ожидания и дисперсии очень просты, но могут оказаться совсем непригодными, если при численной оценке математического ожидания и дисперсии возникнут дополнительные погрешности, особенно в том случае, если вычислительное устройство работает с небольшим числом разрядов. При большом объеме выборки N начиная с некоторого номера отсчета порядок, т.е. число разрядов суммы, будет существенно больше числа разрядов слагаемых. В результате произойдет потеря точности. Это происходит из-за того, что при суммировании чисел происходит выравнивание порядков. Т.е. мантисса числа с меньшим порядком сдвигается вправо на количество разрядов, равное разности порядков суммируемых чисел. Если эта разность велика, то часть или все разряды слагаемого могут выйти за разрядную сетку вычислителя. Поэтому при цифровой обработке, чтобы избежать этого эффекта, можно использовать более сложные алгоритмы вычисления. Этот алгоритм используется, если $N = 2^r$. Введем новую функцию $f(x, y) = \dfrac{x+y}{2}$.

Имеем массив чисел $x_1, x_2, \ldots, x_N$.

Алгоритм вычисления математического ожидания имеет вид:

1) $k = 0$;
2) $S_k = f(x_{2k}, x_{2k+1}), \qquad k = k+1$;
3) если $k < N/2$ идти на (2);
4) $N_2 = N/2$;
5) $N_2 = N_2/2$, $k = 0$;
6) $S_k = f(S_{2k}, S_{2k+1}), \qquad k = k+1$;
7) если $k \le N_2$ идти к (6);
8) если $N_2 > 1$ идти к (5);
9) $m_x = S_0$.

В этом алгоритме все время складываются два числа, порядки которых близки, поэтому потери точности не происходит.

Кроме того, есть особенности при вычислении дисперсии. Для вычисления дисперсии наиболее просто с практической точки зрения использовать формулу

$$\sigma_x^2 = \frac{1}{N-1}\left\{\sum_{n=0}^{N-1} x_n^2 - \frac{1}{N}\left(\sum_{n=0}^{N-1} x_n\right)^2\right\}.$$

В этой формуле внутри фигурных скобок есть две суммы, которые называют достаточными статистиками. Для их получения и для расчёта дисперсии не нужно в практических измерениях помнить все отсчёты сигнала. Достаточно иметь всего две переменные для накопления суммы всех чисел и суммы их квадратов. Причём это можно делать в одном цикле. Однако слагаемые в скобках при вычислении могут оказаться близкими по порядку. При небольшом числе значащих цифр и при $N \to \infty$ оба числа будут одного порядка и при вычислении число, которое представляет дисперсию, будет содержать мало значащих цифр. Таким образом, в этих вычислениях происходит потеря точности.

Исходная формула менее чувствительна к этому эффекту, но вычисления по ней требуют большего объема памяти и времени. Приходится хранить в памяти полностью весь массив отсчётов сигнала и выполнять расчёты с помощью организации двух циклов.

Оценка законов распределения

Существуют два типа законов распределения – закон распределения вероятностей и закон распределения плотности вероятности.

Первоначально рассмотрим оценки функции плотности вероятности (ФПВ). Для их оценки используются параметрические и непараметрические методы.

Непараметрические методы

Эти методы используются, когда нет никакой априорной информации о типе закона распределения. Но некоторые предположения должны быть: например, что он непрерывный. Существуют разновидности этих методов оценки:

1. Гистограммный метод.
2. Метод Парзена.
3. Метод разложения по базисным функциям.
4. Метод полигонов Смирнова.
5. Метод К-ближайших соседей.

Гистограммный метод получил самое большое распространение. Применяется гистограммный метод для априорно непрерывных законов распределения плотности вероятности, отличных от нуля, на всем рассматриваемом множестве значений аргумента x и в случае, когда объем выборки достаточно велик. Область возможных значений сигнала $(x_{\min}, x_{\max})$ разбивается на M непересекающихся одинаковых или различных по размерам подобластей. В одномерном случае это интервалы Δ_{xk}, а в многомерном – параллелепипеды. Затем подсчитывается число отсчётов $\hat{\nu}_k(x)$ исходной выборки объёмом N, попавших в каждую подобласть, и вычисляется оценка плотности вероятности по формуле

$$\hat{w}(x) = \frac{\hat{\nu_k}(x)}{\Delta x k N}.$$

Чаще всего гистограммный метод применяется в одномерном случае. Причём размеры интервалов задают одинаковыми. Тогда график полученной оценки имеет вид, показанный на рис. 2.

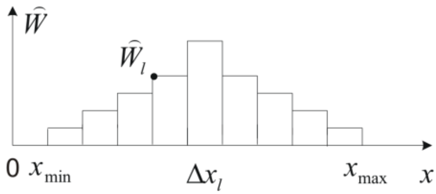

Рисунок 2 – Гистограмма

Это ступенчатая функция, площадь под которой равна 1:

$$S = \hat{W}_l \Delta x_l = 1.$$

Этот признак используется для проверки правильности процесса оценки.

В случае построения гистограммы необходимо правильно выбирать число подобластей или их величину. На первый взгляд кажется, что чем больше интервалов, тем лучше. Но это не так. Число интервалов сверху ограничивают две причины.

Первая – оценки w_n – это случайные величины. Величина ν_n может принимать только целые значения, и при малых ν_n относительный разброс очень велик.

Вторая – квантованный характер данных. Ясно, что бессмысленно брать N больше, чем число уровней квантования, но и при меньших N возможны ситуации, резко искажающие истинное поведение плотности вероятности.

При обработке цифровых сигналов количество интервалов целесообразно выбирать кратным 2^r в пределах (8…16) При этом границы интервалов совпадут с некоторыми уровнями квантования.

Для выбора количества интервалов M есть практические рекомендации. Обычно их число 10, 12, 15 при объёме выборки не менее 100.

Достоинством гистограммного метода являются простота оценивания и ясный физический смысл оценки.

Недостатки гистограммной оценки: если N велико, а количество интервалов неизменно, то оценка сходится не самой ФПВ, а к интегралу от нее: $W_1 \to \int\limits_{\Delta x} W(x)dx$.

Т.е. эта оценка несостоятельная, однако есть методы, позволяющие обойти эту неприятность.

Улучшение гистограммного метода

Для того чтобы оценка сошлась к точному значению ФПВ, подобласти должны выбираться и изменяться таким образом, чтобы выполнялись три условия:

1) $\lim\limits_{N\to\infty} \Delta V = 0$;
2) $\lim\limits_{N\to\infty} \nu_n = \infty$;
3) $\lim\limits_{N\to\infty} \frac{\nu_n}{N} = 0$.

Первое условие обеспечивает сходимость пространственно усреднённого отношения $P/\Delta V$ к плотности вероятности при однородном сокращении областей и при непрерывном характере закона распределения. Второе условие имеет смысл, если во всей области определения $w(x) \neq 0$, и обеспечивает сходимость по вероятности. Третье условие необходимо, если оценка вообще может сойтись к точному значению.

Существует два способа получения последовательности подобластей, удовлетворяющих этим условиям.

Первый заключается в сжатии начальных областей таким образом, чтобы их объём как функция от N был $V_n \equiv 1/\sqrt{N}$. Это метод Парзеновского окна.

Во втором случае размер подобластей выбирается так, чтобы число отсчётов, попавших в каждую из подобластей, было пропорционально $\sqrt{N}$. Т.е. объём изменяется так, чтобы охватить ν_n соседей. Это метод оценки по ν_n ближайшим соседям.

Непараметрическая оценка закона распределения вероятности

От недостатков оценки плотности распределения вероятности свободна оценка интегрального закона распределения:

$$\hat{P}_m\{x_k \le x_{мин} + m\Delta_x\} = \frac{n_m}{N},$$

$$\hat{P}_0\{x \le x_{мин}\} = 0,$$

$$\hat{P}_N\{x \le x_{макс}\} = 1.$$

В результате получается ступенчатая функция, показанная на рис. 3.

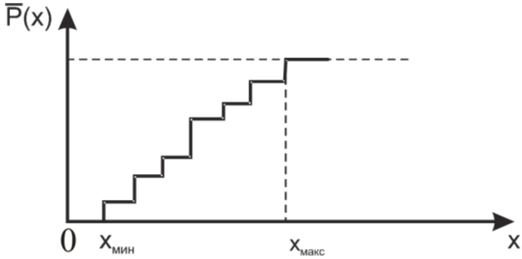

Рисунок 3 – Оценка закона распределения вероятности

В случае аналогового сигнала оценка $\hat{P}_n$ стремится по вероятности к P_n при $N \to \infty$. При этом количество ступенек также должно стремиться к бесконечности, а величина каждой ступеньки – к нулю. В результате ступенчатая кривая стремится к непрерывной.

Для цифрового сигнала это невозможно из-за квантованности, вследствие чего количество ступенек не может быть более 2^r, что ограничивает точность оценки.

Параметрический метод оценки распределений. Аппроксимация экспериментальных распределений

Сама по себе экспериментально полученная оценка кривой плотности распределения вероятности представляет мало ценности. В подавляющем большинстве случаев необходимо найти аналитическую запись закона распределения. Для решения этой задачи можно воспользоваться несколькими методами.

1. Можно пользоваться своим предшествующим опытом, интуицией. Строят гистограмму и затем её сглаживают, как показано на рис. 4.

По виду сглаженной кривой судят о типе распределения. Самый существенный недостаток этого метода – разные люди придут к разным выводам о типе распределения.

2. Выбор типа распределения на основе изучения гистограмм и физической природы явления. Этот метод применим в очень редких случаях. Очень часто трудно сопоставить физику процесса и его вероятностные характеристики.

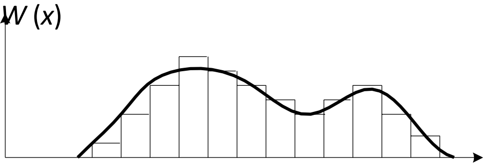

Рисунок 4 –Сглаженная гистограмма

3. Очень часто применяется метод моментов. Этот метод позволяет устранить субъективизм, формализовать процедуру подбора распределений, автоматизировать всю процедуру анализа.

Отметим, что начальным моментом k-го порядка является величина

$$\mu_k = \int_{-\infty}^{\infty} x^k w(x) dx .$$

А центральным моментом k-го порядка называют

$$M_k = \int_{-\infty}^{\infty} (x-\mu_1)^k dx .$$

Моменты являются величинами, которые характеризуют форму распределения. Если существуют все моменты, то можно утверждать, что они полностью и однозначно определяют вид функции распределения. Это утверждение справедливо для подавляющего большинства видов распределений. Т.е. можно утверждать, что знание моментов эквивалентно знанию функции распределения.

В большинстве случаев для удовлетворительной подгонки экспериментальной функции распределения достаточно знать всего четыре первых центральных момента. И даже не сами моменты, а коэффициенты, связанные с ними функционально:

коэффициент асимметрии $\beta_1 = \frac{M_3^2}{M_2^3}$

и коэффициент эксцесса $\beta_2 = \frac{M_4}{M_2^2}$.

На практике мы знаем не точные значения моментов и этих коэффициентов, а их оценки $\hat{\beta}_1$, $\hat{\beta}_2$, полученные через оценки моментов

$$\hat{\mu}_1 = \frac{1}{N}\sum_{i=1}^{N} x_i;$$

$$\hat{M}_k = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \hat{\mu}_1\right)^k.$$

С увеличением объёма выборки точность оценки, как правило, увеличивается.

Оказывается, что если нанести точки на график в осях (β_1, β_2), то для каждого распределения на этой плоскости существует своё место, т.е. точка, линия или целая область. Эту плоскость называют плоскостью Пирсона (рис. 5).

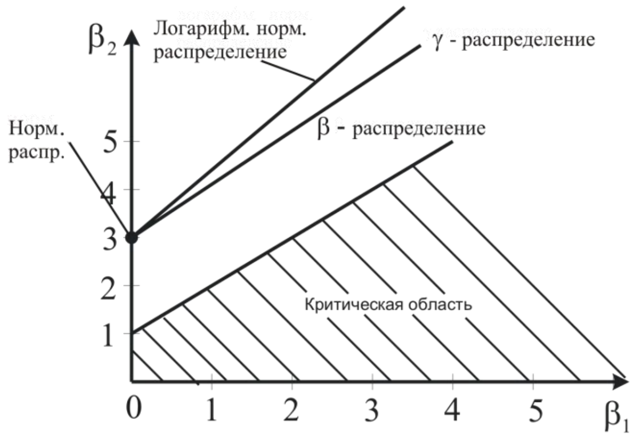

Рисунок 5 – Плоскость Пирсона

На эту плоскость наносят оценки $\hat{\beta}_1$, $\hat{\beta}_2$ и смотрят, к какому распределению ближе всего ложится эта точка. При этом в некоторых местах этой плоскости возможен неоднозначный выбор типа распределения.

Когда тип распределения определён, процедура подбора ещё не закончена, так как кроме типа распределения необходимо найти его параметры. Параметры распределения желательно искать, если это возможно, методом максимального правдоподобия. Этот метод наиболее эффективен. Напомним, что в этом методе в качестве оценки параметра выбирают такую величину, которая максимизирует совместную функцию распределения всех элементов выборки путём решения уравнения правдоподобия. Если оценивается несколько параметров, то получают систему уравнений правдоподобия. Метод максимального правдоподобия даёт наилучшую оценку.

В случае когда методом максимального правдоподобия не удаётся воспользоваться из-за математических сложностей решения системы уравнений правдоподобия, применяют метод моментов.

При этом оценки моментов, полученные экспериментальным путём, приравнивают к теоретически полученным выражениям для моментов распределения, выбранного на первом этапе. При этом используют столько моментов, сколько неизвестных параметров распределения требуется определить. Затем решают полученную систему уравнений. Часто это решение бывает непростым.

Оценка ковариации

Автокорреляционной функцией (АКФ) называют

$$R_{xx}(\tau) = \operatorname{cov}[x] = \int_{-\infty}^{\infty} [x(t) - m_x][x(t-\tau) - m_x] W(x,\tau)\,dx .$$

Для упрощения рассуждений, но без потери общности будем считать, что $x(t)$ – это стационарный случайный процесс с нулевым математическим ожиданием $m_x = 0$. Тогда получаем автоковариационную функцию

$$B_{xx}(\tau) = \int_{-\infty}^{\infty} x(t)x(t-\tau)W(x,\tau)\,dx = R_{xx}(\tau),$$

которая совпадает с АКФ.

Используя понятие выборочного среднего, можно записать оценку автокорреляционной последовательности:

$$\hat{B}_{xx} = \frac{1}{N}\sum_{k=0}^{N-|m|-1} x_k\, x^{*}_{k+m},\ m = 0, \pm 1, \pm 2, \ldots \text{ и } |m| < N .$$

Математическое ожидание этой оценки не совпадает с точным значением автокорреляционной функции

$$E\left[\hat{B}_{xx}(m)\right] = \frac{N - |m|}{N} B_{xx}(m),$$

т.е. оценка автоковариационной функции является смещённой. Но её можно скорректировать так же, как и оценку дисперсии

$$\hat{B}'_{xx} = \frac{1}{N-|m|}\sum_{k=0}^{N-|m|-1} x_k x^*_{k+m} = \frac{N}{N-|m|}\hat{B}_{xx}(m).$$

Это выражение даёт приемлемую оценку автоковариационной функции для различных законов распределения исходной последовательности отсчётов сигнала. Она является несмещённой и её дисперсия равна:

$$\mathrm{var}\left[\hat{B}'_{xx}(m)\right] \cong \frac{N}{\left(N-|m|\right)^2}\sum_{r=-\infty}^{\infty}\left[\mathrm{B}^2_{\mathrm{xx}}(r) + B_{xx}(r+m)B_{xx}(r-m)\right].$$

Эта формула справедлива при $\mathrm{N} >> \mathrm{m}$.

Отсюда видно, что

$$\lim_{\mathrm{N}\to\infty}\left\{\mathrm{var}\left[\hat{B}'_{xx}(m)\right]\right\} \to 0.$$

Т.е. оценка автоковариации, как и АКФ, является состоятельной.

Цифровые методы спектрального анализа

Основной целью спектрального анализа является оценивание спектральной плотности мощности (СПМ) дискретизированного процесса и обнаружение присутствия в нём в течение определённого интервала времени периодического сигнала с определёнными параметрами. Под СПМ понимают преобразование Фурье от АКФ:

$$G_{xx}(\omega) = \int_{-\infty}^{\infty} B_{xx}(\tau)\exp(-j\omega\tau)d\tau.$$

Другим определением СПМ является

$$G_{xx}(\omega) = \lim_{T\to\infty}\left\{\frac{1}{T}\left|\int_{-\infty}^{\infty} x(t)\exp(-j\omega t)\,dt\right|^2\right\}.$$

Это предел модуля квадрата преобразования Фурье от обрабатываемого сигнала.

Эти выражения являются основой для разных методов оценки СПМ.

Обработка дискретизированного процесса обычно производится последовательно во времени, причём одновременно обрабатывается выборка $x_1, x_2, \ldots, x_N$ из N отсчётов, полученных из сигнала $x(t)$ с шагом Δt. Интервал $T_н = N\Delta t$ называют длиной реализации или интервалом наблюдения.

Коррелограммный метод оценки СПМ

В качестве оценки СПМ обычно рассматривают преобразование Фурье смещённой оценки автокорреляции:

$$\hat G(n) = \sum_{m=-(L-1)}^{L-1} \hat B_{xx}(m)\exp\left(j\frac{2\pi}{N}nm\right), \qquad n = 0, \pm1, \pm 2, \dots, \pm L; \quad L \approx 0{,}1N .$$

Оценка автокорреляции является состоятельной, в то время как её преобразование Фурье – это несостоятельная оценка. Определим смещение оценки

$$E\left\{\sum_{m=-(L-1)}^{L-1} \hat B_{xx}(m)\exp\left(j\frac{2\pi}{N}nm\right)\right\} = \sum_{m=-(L-1)}^{L-1} E\left[\hat B_{xx}(m)\exp\left(j\frac{2\pi}{N}nm\right)\right] =$$

$$= \sum_{m=-(L-1)}^{L-1} \frac{N-|m|}{N} B_{xx}(m)\exp\left(j\frac{2\pi}{N}nm\right).$$

Т.е. оценка асимптотически не смещена, так как при увеличении N множитель под знаком суммы стремится к 1. Множитель под знаком суммы можно рассматривать как весовую функцию треугольного вида:

$$w_B(m) = \begin{cases} \dfrac{N-|\mu|}{N} & |m| < N \\ 0 & |m| \ge N \end{cases}.$$

Это так называемое окно Бартлетта. Следовательно математическое ожидание оценки СПМ, полученной коррелограммным методом, является преобразованием Фурье АКФ с весовой функцией Бартлетта.

Периодограммный метод оценки СПМ

В этом случае за оценку СПМ принимают

$$\hat G(\omega) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x_n e^{-j\omega n\Delta t}\right|^2 .$$

Оценим её характеристики:

$$E\left[\hat G(n)\right] = G(n) * W_B(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W_B(\omega-\nu)G(\nu)\,d\nu ,$$

где $*$ - знак свёртки, $W_B(\omega) = \frac{1}{N}\left[\frac{\sin(N\omega/2)}{\sin(\omega/2)}\right]^2$.

Это свёртка истинной СПМ со спектром весовой функции Бартлетта, т.е. сглаженная оценка. Поэтому периодограмма является величиной смещённой. Причём сглаживание приводит к ухудшению разрешающей способности, т.е. способности различать близко расположенные узкополосные составляющие спектра. Разрешение периодограммы равно

$$\Delta\omega = 0{,}89\left(2\pi/N\right).$$

Поиск дисперсии периодограммы в общем случае очень сложен. В частном случае белого нормального шума дисперсия периодограммы равна

$$\operatorname{var}\left[\hat{G}(\omega)\right] = G^2(\omega)\left\{1 + \left[\frac{\sin(\omega N\Delta t)}{N\sin[\omega\Delta t]}\right]\right\}.$$

Отсюда следует, что оценка СПМ периодограммным методом несостоятельна. При увеличении N дисперсия оценки СПМ становится пропорциональной квадрату точного значения СПМ и сильно флюктуирует около истинного значения. Во многих случаях несостоятельность оценки СПМ делает невозможным её использование в практических задачах. Однако лёгкость её получения с помощью БПФ заставляет искать пути снижения дисперсии оценки.

Причина несостоятельности оценки СПМ состоит в том, что оценивается не числовой параметр случайного процесса, а функция. При оценке параметра увеличение объёма выборки приводит к увеличению количества информации, используемого в этой оценке. При оценке СПМ происходит не уточнение уже имеющихся значений отсчётов СПМ, а вычисление новых отсчётов.

Сглаживание спектральных оценок

Уменьшить дисперсию оценки СПМ позволяют различные методы сглаживания. Широко известны три таких метода.

Метод Блэкмана – Тьюки

Этот метод заключается в сглаживании оценки периодограммы путём свёртки её с подходящим спектральным окном:

$$\bar{G}(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\hat{G}(\Theta)W(\omega-\Theta)\,d\Theta$$

или умножением корреляционной функции при вычислении преобразования Фурье на весовую функцию

$$\bar{G}(\omega) = \sum_{m=-(N/2)-1}^{(N/2)} \hat{B}_{xx}(m)\,w(m)\,e^{-j\omega m\Delta t}.$$

Эти преобразования можно продемонстрировать на примере прямоугольного окна (Окна Пугачёва), показанного на рис. 6.

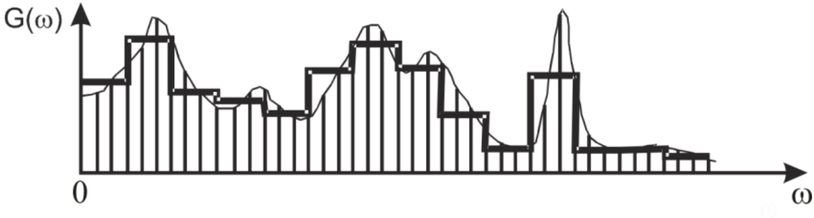

Рисунок 6 – Сглаживание периодограммы окном Пугачёва

Определим свойства этой оценки. Математическое ожидание

$$E\left[\tilde{G}(\omega)\right] = \frac{1}{2\pi} G(\omega) * W(\omega).$$

Дисперсия оценки

$$\operatorname{var}\left[\tilde{G}(\omega)\right] = G(\omega)\frac{1}{N}\sum_{k=-M}^{M} w^2(k).$$

Отсюда видно, что при выборе ширины окна M необходимо принимать компромиссное решение. Для получения малого смещения необходимо, чтобы M было достаточно большим для снижения ширины главного лепестка спектра весовой функции. С другой стороны, M должно быть достаточно малым, чтобы сумма в последнем выражении была маленькой.

Разрешение по частоте в этом методе зависит от вида используемого весового окна (или весовой функции).

Периодограмма Бартлетта

Вся последовательность исходных данных

$$\{x_k\}, \quad k = 0,1,2,\ldots,(N-1)$$

делится на K отрезков по M отсчётов так, чтобы

$$N = KM.$$

Затем для каждого отдельного отрезка вычисляется периодограмма

$$\tilde{G}_M^{(i)}(\omega) = \frac{1}{M}\left|\sum_{n=0}^{M-1} x_{(i)}(n) e^{-j\omega n\Delta t}\right|^2.$$

Таких периодограмм всего K. Далее производится усреднение

$$\tilde{G}(\omega) = \frac{1}{K}\sum_{i=1}^{K} \tilde{G}_M^{(i)}(\omega).$$

Определим смещение

$$E\left[\hat{G}(\omega)\right] = \frac{1}{2\pi} G(\omega) * W_B(\omega).$$

Такая оценка является смещённой, но при этом дисперсия оценки уменьшается в K раз:

$$\mathrm{var}\left\{\hat{G}(\omega)\right\} = \mathrm{var}\left\{\frac{1}{K}\sum_{i=1}^{K}\hat{G}_M^{(i)}(\omega)\right\} = \frac{\sum_{i=1}^{K}\mathrm{var}\left\{\hat{G}_M^{(i)}(\omega)\right\}}{K^2} = \frac{\mathrm{var}\left\{\hat{G}_M^{(i)}(\omega)\right\}}{K}.$$

Разрешение по частоте

$$\Delta\omega = 0{,}89K\left(2\pi/N\right).$$

Периодограмма Уэлча

Это компромисс между двумя предыдущими методами. Ещё он называется модифицированной периодограммой. Здесь также происходит деление всего массива исходных отсчётов на отдельные отрезки (сегменты) по M отсчётов. Причём эти сегменты смещены по времени на D отсчётов. Перекрытие необходимо для увеличения общего числа таких сегментов. Тогда i-й сегмент можно записать как

$$x_i(n) = x(n+iD),\ n = 0,1,\dots,M-1.$$

Следовательно, перекрытие между $x_i(n)$ и $x_{i+1}(n)$ равно $M-D$ отсчётов. Если K сегментов покрывают все N исходных отсчётов, то

$$N = M + D(K-1).$$

Если перекрытия нет, т.е. $D = M$, имеем $K = N/M$, как в методе Бартлетта. При перекрытии 50 % ($D = M/2$) можно сформировать $K = 2M/(M-1)$ сегментов, а следовательно, сохранить разрешение, как и в методе Бартлетта, но при удвоении числа усреднённых периодограмм, и за счёт этого уменьшить дисперсию оценки. С другой стороны, при 50%–м перекрытии можно сформировать $K = N/M - 1$ последовательности длиной $2M$ и тем самым повысить разрешение, сохраняя то же значение дисперсии, как и для метода Бартлетта. Т.е. в этом методе возможен компромисс между дисперсией и разрешением.

СПМ для каждого сегмента вычисляется с использованием весовой функции $w(n)$:

$$\hat{G}_M^{(i)}(\omega) = \frac{1}{MU}\left|\sum_{n=0}^{M-1} x_n w(n) e^{-j\omega n\Delta t}\right|^2,$$

где $U = \dfrac{1}{M}\sum\limits_{n=0}^{M-1} w^2(n)$ - нормирующий множитель.

Окончательная оценка вычисляется путём усреднения спектров каждого отрезка:

$$\hat{G}(\omega) = \frac{1}{K}\sum_{i=1}^{K}\hat{G}_M^{(i)}(\omega).$$

Значит, можем уменьшить дисперсию в K раз. Однако при наличии перекрытия дисперсия уменьшается меньше. Так, при использовании окна Бартлетта и 50%–м перекрытии

$$\mathrm{var}\left\{\hat{G}(\omega)\right\} = \frac{9}{8}\,\frac{\mathrm{var}\left\{\hat{G}_M^{(i)}(\omega)\right\}}{K}.$$

Оценим смещение:

$$E\left[\hat{G}(\omega)\right] = \frac{1}{MU}\,G(\omega)\left|W(\omega)\right|^2.$$

Таким образом, метод Уэлча даёт асимптотически несмещённые оценки СПМ.

Сравнение характеристик методов оценки СПМ

Видно, что имеется компромисс между разрешающей способностью по частоте и дисперсией оценки. Для сравнения характеристик этих методов используют два критерия. Первый из них представляет собой нормированную дисперсию оценки:

$$\alpha = \mathrm{var}\left[G(\omega)\right] \big/ E^2\left[G(\omega)\right].$$

Он характеризует меру случайного разброса оценки. Т.е. это мера статистической устойчивости оценки СПМ.

Второй критерий – это произведение устойчивости на разрешение:

$$\mu = \alpha\Delta\omega.$$

Этот критерий позволяет определить разрешение СПМ, получаемой по конечной последовательности отсчётов реализации с учётом случайного характера сигнала. Качество спектральной оценки тем выше, чем меньше μ.

В табл. 1 приведены эти параметры и разрешение по частоте для рассмотренных методов сглаживания оценки СПМ.

Из таблицы видно, что каждый метод имеет приблизительно одно и то же качество оценки, обратно пропорциональное длине последовательности данных. Хотя каждый из методов отличается по дисперсии и разрешению, общая характеристика ограничена количеством доступных данных.

Таблица 1

Метод	α	Δω	μ
Периодограмма	1	$0{,}89(2\pi/N)$	$0{,}89(2\pi/N)$
Бартлетта	$1/K$	$0{,}89(2\pi/N)$	$0{,}89(2\pi/N)$
Блэкмана – Тьюки	$(2/3)(M/N)$	$0{,}64K(2\pi/M)$	$0{,}43(2\pi/N)$
Уэлча (50%–й)	$(9/8)(1/K)$	$1{,}282(2\pi/M)$	$0{,}72(2\pi/N)$

Во всех рассмотренных методах сглаживания оценки СПМ возникает смещение оценки из-за ухудшения разрешающей способности по частоте, но происходит снижение дисперсии. Поэтому здесь необходим компромисс между смещением и дисперсией, решаемый в каждой задаче конкретно с учётом априорных сведений. Например, если мы знаем или ожидаем, что спектр может иметь узкий пик и его важно выделить, то нужно выбрать M достаточно большим для получения необходимой разрешающей способности. А для получения требуемой степени снижения дисперсии оценки необходимо выбирать достаточно большое число K. Это приводит к необходимости существенного увеличения общего объёма выборки N. Но во всех реальных задачах невозможно увеличивать объем выборки бесконечно. Не надо забывать, что для снижения дисперсии оценки СПМ усредняемые значения должны быть независимыми.

Индивидуальное задание для расчёта

Разработать в системе программирования Matlab программу для оценки параметров и характеристик реализации случайного процесса и выполнить все расчёты, указанные ниже.

Описание исходных данных

Исходный массив данных, представляющих собой реализацию из 2048 отсчетов многомерного случайного сигнала, записан в файлах «EEG1(2,3,4).txt» текстового формата. Количество колонок массива соответствует числу компонент случайного процесса (16), количество строк – числу отсчетов (2048). Интервал дискретизации, использованный при регистрации сигнала, равен 1 мс.

Внимание! Для упрощения программы файлы с исходными данными должны быть сохранены в той же папке, что и текст программы.

Выбрать задание на исследование в соответствии с табл. 2, где номер варианта соответствует номеру студента в списке для 16–й группы и номеру, увеличенному на 20 для 17–й группы. Номер канала – номер столбца данных, который требуется обработать.

Таблица 2

Вариант	Номера каналов	Спектральное окно	Размер интервала	Номер файла
1	1,2,3	Треугольное	128	1
2	4,5,6	Бартлетта	256	2
3	7,8,9	Ханна	512	3
4	10,11,12	Хэмминга	128	4
5	13,14,15	Блэкмана	256	1
6	16,1,2	Кайзера	512	2
7	3,4,5	Чебышева	128	3
8	6,7,8	Тьюки	256	4
9	9,10,11	Треугольное	512	1
10	12,13,14	Бартлетта	128	2
11	15,16,1	Ханна	256	3
12	2,3,4	Хэмминга	512	4
13	5,6,7	Блэкмана	128	1
14	8,9,10	Кайзера	256	2
15	11,12,13	Чебышева	512	3
16	14,15,16	Тьюки	128	4
17	1,3,5	Треугольное	256	1
18	7,9,11	Бартлетта	512	2
19	13,15,1	Ханна	128	3
20	3,5,7	Хэмминга	256	4
21	9,11,13	Блэкмана	512	3
22	15,1,3	Кайзера	128	4
23	5,7,9	Чебышева	256	1
24	11,13,15	Тьюки	512	2
25	2,4,6	Треугольное	128	3
26	8,10,12	Бартлетта	256	4
27	14,16,2	Ханна	512	1
28	4,6,8	Хэмминга	128	2
29	10,12,14	Блэкмана	256	3

Окончание табл. 2

Вариант	Номера каналов	Спектральное окно	Размер интервала	Номер файла
30	16,2,4	Кайзера	512	4
31	6,8,10	Чебышева	128	1
32	12,14,16	Тьюки	256	2
33	1,4,7	Треугольное	512	3
34	2,5,8	Бартлетта	128	4
35	3,6,9	Ханна	256	1
36	4,7,10	Хэмминга	512	2
37	5,8,11	Блэкмана	128	3
38	6,9,12	Кайзера	256	4
39	7,10,13	Чебышева	512	1
40	8,11,14	Тьюки	128	2

Предварительная обработка

Построить графики зависимости обрабатываемых данных от времени, т.е. от номера отсчёта.

Задачи для решения

Задача 1. Оценка статистических характеристик реализации случайного процесса.

Выполнить оценку математического ожидания и дисперсии, заданных в соответствии с номером варианта компонент реализации СП. Оценить их вариативность (дисперсию), выполнив расчеты для разных временных интервалов. Сделать выводы.

Задача 2. Оценка плотности распределения реализации случайного процесса.

Выбрать количество интервалов, рассчитать и построить гистограмму распределения для заданных компонент реализации СП.

Задача 3. Оценка корреляционных характеристик реализации случайного процесса.

Оценить и построить корреляционные функции заданных в соответствии с номером варианта компонент СП.

Оценить и построить взаимные корреляционные функции компонент СП.

Задача 4. Оценка спектральных характеристик реализации случайного процесса.

Исследовать спектральные характеристики (СПМ) случайного процесса методами классического спектрального оценивания.

Четные варианты: модифицированным периодограммным методом, нечетные варианты: методом Уэлча.

Длительности интервалов (сегментов сигнала) выбрать из табл. 2.

Для снижения дисперсии оценки использовать выделяющие функции или сглаживание спектральными окнами, выбранными в соответствии с вариантом.

Сравнить оценки СПМ, выполненные без использования и с использованием выделяющих функций, а также в зависимости от размера выделяющей функции.

Требования к выполнению курсовой работы

Данные для выполнения задачи следует выбирать из соответствующей таблицы согласно своему варианту, указанному преподавателем.

1. Все задания должны быть выполнены самостоятельно после изучения соответствующего раздела курса.
2. Задания выполняются на стандартных листах писчей бумаги формата А4 (297х210 мм).
3. Пояснительная записка к курсовой работе выполняется либо чернилами, либо печатным способом, записи ведутся только на одной стороне листа.
4. Графическая часть задания выполняется в масштабе с использованием чертежного инструмента либо печатным способом.
5. Работы, выполненные с нарушением этих указаний, не рассматриваются.

Содержание пояснительной записки

1. Основные теоретические сведения по всем оцениваемым параметрам (точное определение оцениваемой величины, формула для оценки, свойства используемой оценки).
2. Результаты расчётов в графической либо численной формах (в зависимости от типа характеристик) с пояснениями и выводами.
3. Полный текст программы с комментариями.
4. Для защиты также должна быть представлена электронная версия программы.

Библиографический список

1. Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения. М.: Мир, 1990. 584 с.
2. Сергиенко А.Б. Цифровая обработка сигналов. СПб.: Питер, 2002.
3. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи. М.: Радио и связь, 2000.
4. Шахтарин Б.И., Ковригин В.А. Методы спектрального оценивания случайных процессов. М.: Гелиос АРВ, 2005. 248 с.

Оглавление