

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ И ИЗОБРАЖЕНИЙ. ВЕРОЯТНОСТНЫЕ МОДЕЛИ

Рекомендовано редакционно-издательским советом федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» в качестве учебного пособия для обучающихся по основной образовательной программе высшего образования по специальности 10.05.03 Информационная безопасность автоматизированных систем

САМАРА

Издательство Самарского университета
2024

УДК 621.391.8(075)+004.932(075)

ББК 3811я7

Т338

Рецензенты: д-р техн. наук, проф. В. Г. Карташевский,
д-р физ.-мат. наук, проф. С. Я. Новиков

Т338 Теоретические основы цифровой обработки сигналов и изображений. Вероятностные модели: учебное пособие / *В.В. Сергеев, М.А. Чичева, Р.Р. Юзькив, А.Ю. Денисова.* – Самара: Издательство Самарского университета, 2024. – 84 с.: ил.

ISBN 978-5-7883-2043-4

В пособии рассмотрены основы теории цифровой обработки сигналов. Состоит из трех разделов, посвященных математическим моделям одномерных и двумерных случайных сигналов с дискретным аргументом, а также задачам фильтрации и восстановления таких сигналов.

Предназначено для обучающихся по специальности «Информационная безопасность автоматизированных систем», по направлению подготовки «Прикладная математика и информатика» и другим информационно-технологическим направлениям высшего образования, а также может быть полезно для аспирантов и научных работников.

УДК 621.391.8(075)+004.932(075)

ББК 3811я7

ISBN 978-5-7883-2043-4

© Самарский университет, 2024

ОГЛАВЛЕНИЕ

Введение	4
1 Математические модели одномерных случайных сигналов	5
1.1 Основные характеристики случайных процессов	5
1.2 Случайные последовательности и их характеристики	9
1.3 Прохождение случайных последовательностей через ЛИС-системы	13
1.4 Факторизация энергетического спектра	16
2 Одномерная линейная фильтрация и восстановление сигналов	25
2.1 Восстановление дискретного сигнала ЛИС-системой	25
2.2 Оптимальное линейное восстановление	31
2.3 Реализация оптимального фильтра обработкой «в прямом и обратном времени»	40
2.4 Реализация оптимального фильтра при помощи ДПФ	46
2.5 Восстановление сигнала физически-реализуемым БИХ фильтром	50
2.6 Восстановление сигнала КИХ-фильтром	59
3 Двумерные случайные сигналы и системы	63
3.1 Двумерные случайные последовательности, их характеристики и прохождение через ЛИС-системы	63
3.2 Двумерная оптимальная линейная фильтрация	67
3.3 Двумерные линейные субоптимальные КИХ-фильтры	78
Библиографический список	83

ВВЕДЕНИЕ

Учебное пособие «Теоретические основы цифровой обработки сигналов и изображений. Вероятностные модели» предназначено для студентов, обучающихся по специальности «Информационная безопасность автоматизированных систем», направлению «Прикладная математика и информатика» и другим информационно-технологическим направлениям высшего образования, а также аспирантов и научных работников.

Учебное пособие состоит из трех разделов, посвященных математическим моделям одномерных и двумерных случайных сигналов с дискретным аргументом, а также задачам фильтрации и восстановления таких сигналов.

Каждый раздел содержит примеры практических заданий, выполнение которых поможет студентам лучше усвоить теоретический материал.

1 МАТЕМАТИЧЕСКИЕ МОДЕЛИ ОДНОМЕРНЫХ СЛУЧАЙНЫХ СИГНАЛОВ

1.1 Основные характеристики случайных процессов

В отличие от детерминированных процессов, течение которых определено однозначно, *случайный процесс (сигнал)* представляет изменения физической системы во времени или в пространстве, которые заранее в точности предсказать невозможно.

Понятие случайного процесса хорошо знакомо. Каждый раз, когда проводится эксперимент (опыт), итогом его является функция, определенная на интервале времени, а не какое-либо одно число. Если f – функция одной переменной, то говорят о случайном процессе, если f – функция двух или большего числа переменных, то говорят о случайном поле.

Аргумент функции f может быть непрерывным и дискретным. В последнем случае используют термин «случайная последовательность» – одномерная (случайный процесс) или многомерная (случайное поле).

Заметим следующее: *каждая отдельная реализация случайного сигнала является функцией детерминированной*. Поэтому для описания индивидуальных свойств реализаций случайного процесса следует использовать методы, изложенные в учебном пособии [1]. Особенности случайного процесса проявляются при изучении свойств совокупности реализаций (ансамбля). Поскольку этот ансамбль – вероятностный, то и характеристики случайного процесса оказываются вероятностными.

Как известно [2], одномерная *функция распределения вероятностей* значений случайной функции $f(t)$, зависящей от непрерывного аргумента (времени) t

$$P_t(\eta) = P\{f(t) < \eta\} \quad (1.1)$$

связана с одномерной *плотностью вероятностей*

$$p_t(\eta) = \frac{\partial P_t(\eta)}{\partial \eta} . \quad (1.2)$$

Соответственно, r -мерная плотность вероятностей:

$$P_{t_1, t_2, \dots, t_r}(\eta_1, \eta_2, \dots, \eta_r) = \frac{\partial^r P_{t_1, t_2, \dots, t_r}(\eta_1, \eta_2, \dots, \eta_r)}{\partial \eta_1 \partial \eta_2 \dots \partial \eta_r} = \frac{\partial^r P_{\mathbf{t}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}, \quad (1.3)$$

где $\mathbf{t} = (t_1, t_2 \dots t_r)$, $\boldsymbol{\eta} = (\eta_1, \eta_2 \dots \eta_r)$.

Плотность вероятностей удовлетворяет условию нормировки:

$$\int_{-\infty}^{\infty} p_t(\eta) d\eta = 1 \quad (1.4)$$

– в одномерном случае, а в r -мерном случае:

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_{\mathbf{t}}(\boldsymbol{\eta}) d\boldsymbol{\eta} = 1 . \quad (1.5)$$

Последовательность функций

$$p_t(\eta), p_{t_1, t_2}(\eta_1, \eta_2), \dots, p_{t_1, t_2, \dots, t_r}(\eta_1, \eta_2 \dots \eta_r)$$

представляет своеобразную лестницу, поднимаясь по которой, удается все более и более подробно характеризовать случайный процесс. В прикладных задачах часто достаточно знать о случайном процессе меньше, чем дают функции распределения: можно ограничиться *числовыми характеристиками* случайного процесса.

Среди числовых характеристик случайного процесса наиболее важными являются *среднее значение* $\mu_f(t)$, *дисперсия* $\sigma_f^2(t)$ и *автокорреляционная функция (АКФ)* $R_f(t, \tau)$:

$$\mu_f(t) = E\{f(t)\}, \quad \sigma_f^2(t) = E\left\{\left(f(t) - \mu_f(t)\right)^2\right\},$$

$$R_f(t, \tau) = E\left\{\left(f(t) - \mu_f(t)\right)\left(f(\tau) - \mu_f(\tau)\right)\right\},$$

где E – оператор усреднения.

АКФ характеризует степень статистической взаимосвязи случайного процесса при различных аргументах. Очевидно, ее значения зависят не только от степени взаимосвязи, но и от самих значений случайного процесса. Эта зависимость устраняется введением нормировки:

$$\rho_f(t, \tau) = \frac{R_f(t, \tau)}{\sqrt{R_f(t, t)R_f(\tau, \tau)}} = \frac{R_f(t, \tau)}{\sigma_f(t)\sigma_f(\tau)}, \quad |\rho_f(t, \tau)| < 1$$

– эту величину называют *коэффициентом корреляции* между значениями процесса.

Для определения меры статистической зависимости между двумя случайными процессами f и g рассматривают *взаимную корреляционную функцию (ВКФ)*:

$$R_{fg}(t, \tau) = E\left\{\left(f(t) - \mu_f(t)\right)\left(g(\tau) - \mu_g(\tau)\right)\right\}.$$

Если описание случайного процесса не выходит за рамки введенных статистических характеристик, говорят, что оно выполнено в рамках корреляционной теории или на уровне статистик второго порядка.

Случайный процесс $f(t)$ называется *стационарным в узком смысле* (строго), если аналитическое выражение плотности вероятности не зависит от выбора точки начала отсчета времени. Из приведенного определения стационарного процесса следует, что одномерная плотность вероятностей не зависит от времени, а для

числовых характеристик стационарного процесса справедливы следующие свойства:

– среднее значение и дисперсия не зависят от времени:

$$\mu_f(t) = m_f, \quad \sigma_f^2(t) = \sigma_f^2; \quad (1.6)$$

– АКФ зависит только от разности аргументов $t = t' - \tau$

$$R_f(t', \tau) = R_f(t' - \tau) = R_f(t). \quad (1.7)$$

При этом

$$|R_f(t)| \leq R_f(0) = \sigma_f^2, \quad R_f(t) = R_f(-t). \quad (1.8)$$

Кроме того, обычно выполняется условие:

$$R_f(t) \rightarrow 0 \quad \text{при} \quad t \rightarrow \infty. \quad (1.9)$$

Случайные процессы, удовлетворяющие условиям (1.6), (1.7) называют *стационарными в широком смысле* (по А.Я. Хинчину). Случайные процессы, стационарные в узком смысле (строго), являются стационарными в широком смысле, но не наоборот.

Стационарный случайный процесс называется *эргодическим*, если любая его вероятностная характеристика может быть получена из одной достаточно длинной его реализации путем усреднения во времени: среднее во времени равно среднему по ансамблю. На практике, как правило, мы не располагаем множеством реализаций случайного процесса, но имеем возможность наблюдать его в течение длинного промежутка времени T или на большом пространственном интервале. В этом случае выражения для оценок математического ожидания и АКФ выглядят следующим образом:

$$\hat{\mu}_f \approx \frac{1}{T} \int_0^T f(t) dt, \quad (1.10)$$

$$\hat{R}_f(\tau) \approx \frac{1}{T} \int_0^T (f(t) - \mu_f)(f(t+\tau) - \mu_f) dt. \quad (1.11)$$

1.2 Случайные последовательности и их характеристики

Произвольная случайная последовательность $f(n)$, т.е. функция целочисленного аргумента n может быть описана посредством указания тех или иных ее статистических характеристик. В дальнейшем ограничимся рассмотрением статистик второго порядка. Для среднего и дисперсии выражения будут иметь вид:

$$\mu_f(n) = E\{f(n)\}, \quad \sigma_f^2(n) = E\left\{(f(n) - \mu_f(n))^2\right\}, \quad (1.12)$$

АКФ последовательности f и ВКФ последовательностей f и g определяются следующим образом:

$$R_f(k, l) = E\left\{(f(k) - \mu_f(k)) \cdot (f(l) - \mu_f(l))\right\}, \quad (1.13)$$

$$R_{fg}(k, l) = E\left\{(f(k) - \mu_f(k)) \cdot (g(l) - \mu_g(l))\right\}. \quad (1.14)$$

Коэффициент корреляции случайной последовательности:

$$\rho_f(k, l) = \frac{R_f(k, l)}{\sigma_f(k)\sigma_f(l)},$$

при этом во многих практических приложениях важную роль играет коэффициент корреляции между соседними отсчетами $\rho_f = E\{\rho_f(n, n+1)\}$.

Условия стационарности случайной последовательности аналогичны условиям для случайных процессов непрерывного аргумента:

$$\mu_f(n) = \mu_f, \quad \sigma_f^2(n) = \sigma_f^2, \quad R_f(k, l) = R_f(k - l), \quad (1.15)$$

Для корреляционных функций стационарных случайных последовательностей справедливы следующие свойства:

$$R_f(0) = \sigma_f^2, \quad R_f(k) = R_f(-k), \quad R_{gf}(k) = R_{fg}(-k), \quad (1.16)$$

$$\lim_{k \rightarrow \infty} R_f(k) = 0, \quad \lim_{k \rightarrow \infty} R_{fg}(k) = 0. \quad (1.17)$$

Везде далее мы ограничимся рассмотрением именно стационарных случайных последовательностей.

Используя свойство эргодичности применительно к случайной последовательности, можно получить оценки ее числовых характеристик. Действительно, пусть число элементов последовательности $1 \leq n \leq N$, тогда дискретные аналоги выражений (1.10) – (1.11) запишутся следующим образом:

$$\hat{\mu}_f \approx \frac{1}{N} \sum_{k=1}^N f(k), \quad (1.18)$$

$$\hat{R}_f(n) \approx \frac{1}{N-n} \sum_{k=1}^{N-n} (f(k) - \mu_f)(f(k+n) - \mu_f). \quad (1.19)$$

Для одномерной стационарной случайной последовательности $f(n)$ АКФ $R_f(m)$ представляет собой одномерную детерминированную последовательность. Введем преобразование Фурье последовательности $R_f(m)$, которое называется *спектральной плотностью мощности* (энергетическим спектром) последовательности $f(n)$:

$$\Phi_f(e^{i\omega}) = \sum_{n=-\infty}^{\infty} R_f(m) e^{-i\omega n}. \quad (1.20)$$

При этом отсчеты АКФ могут быть вычислены через спектральную плотность $\Phi_f(e^{i\omega})$ с помощью обратного преобразования Фурье:

$$R_f(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_f(e^{i\omega}) e^{i\omega m} d\omega. \quad (1.21)$$

Выражения, аналогичные (1.20) и (1.21), связывают ВКФ и взаимный энергетический спектр.

Отметим некоторые свойства энергетических спектров:

– энергетический спектр $\Phi_f(e^{i\omega})$ – вещественная функция частоты;

– энергетический спектр всегда неотрицателен: $\Phi_f(e^{i\omega}) \geq 0$;

– энергетический и взаимный энергетический спектры обладают свойствами симметрии:

$$\Phi_f(e^{i\omega}) = \Phi_f(e^{-i\omega}), \quad \Phi_{fg}(e^{i\omega}) = \Phi_{gf}(e^{-i\omega}).$$

Иногда энергетическими спектрами называют и z -преобразование корреляционных функций.

Рассмотрим примеры.

Пример 1.1. *Белый шум* (последовательность независимых случайных величин). Его корреляционная функция имеет вид

$$R_f(n) = \sigma_f^2 \delta(n). \quad (1.22)$$

Из (1.20) следует

$$\Phi_f(e^{i\omega}) = \sigma_f^2, \quad -\pi \leq \omega \leq \pi, \quad (1.23)$$

то есть спектральная плотность мощности белого шума постоянна на всех частотах (см. рисунок 1.1)

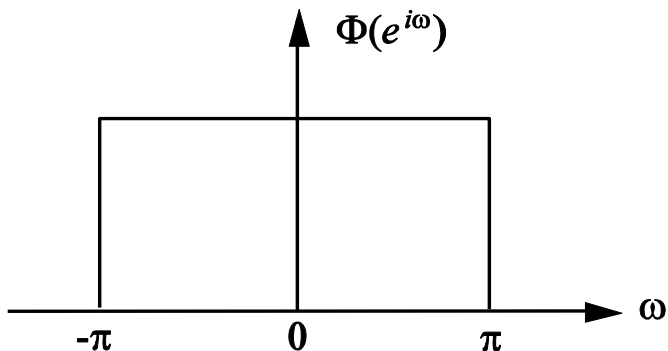


Рис. 1.1. Спектральная плотность мощности последовательности типа «белый шум»

Пример 1.2. Стационарная случайная последовательность с биэкспоненциальной АКФ

$$R_f(n) = \sigma_f^2 \cdot \rho^{|n|} \quad (1.24)$$

имеет энергетический спектр следующего вида (см. рисунок 1.2):

$$\Phi_f(e^{i\omega}) = \frac{\sigma_f^2 \cdot (1 - \rho^2)}{1 + \rho^2 - 2\rho \cos(\omega)}, \quad -\pi \leq \omega \leq \pi,$$

где ρ – коэффициент корреляции между соседними отсчетами последовательности.

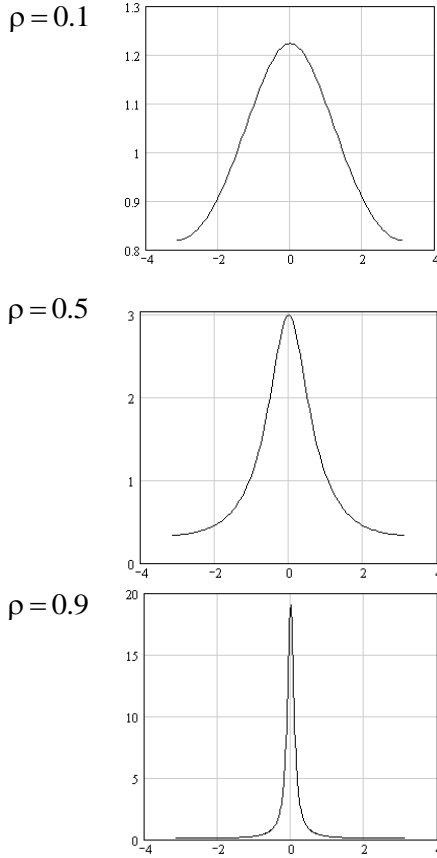


Рис. 1.2. Спектральная плотность мощности случайной последовательности с биекспоненциальной АКФ

1.3 Прохождение случайных последовательностей через ЛИС-системы

Пусть известны характеристики входного сигнала – стационарной случайной последовательности $f(n)$: среднее значение μ_f , АКФ R_f и энергетический спектр Φ_f . Требуется получить

соответствующие характеристики для последовательности $g(n)$ на выходе устойчивой линейной инвариантной к сдвигу (ЛИС) системы с импульсной характеристикой $h(n)$, а также взаимные статистические характеристики входной и выходной последовательностей.

Среднее значение для выходной последовательности с учетом стационарности сигналов и известной формулы свертки определяется следующим образом:

$$\begin{aligned} \mu_g &= E\{g(n)\} = E\left\{\sum_{k=-\infty}^{\infty} h(k)f(n-k)\right\} = \\ &= \sum_{k=-\infty}^{\infty} h(k)E\{f(n-k)\} = \mu_f \sum_{k=-\infty}^{\infty} h(k). \end{aligned} \quad (1.25)$$

Если ЛИС-система описана не импульсной характеристикой, а частотной $H(e^{i\omega})$ или передаточной функцией $H(z)$, то для вычисления среднего значения выходной последовательности можно воспользоваться соотношениями:

$$\mu_g = \mu_f H(e^{i\omega})\Big|_{\omega=0}, \quad \mu_g = \mu_f H(z)\Big|_{z=1}, \quad (1.26)$$

используемыми указанные характеристики системы.

В дальнейшем для сокращения изложения будем полагать $\mu_g = \mu_f = 0$. При невыполнении этого равенства всегда можно учесть математическое ожидание и его преобразование отдельно на основании формул (1.25) и (1.26).

АКФ выходной последовательности

$$R_g(n) = E\left\{\left(g(k) - \mu_g\right)\left(g(k+n) - \mu_g\right)\right\}$$

определяется следующим образом:

$$R_g(n) = \sum_{k=-\infty}^{\infty} \left[\sum_{l=-\infty}^{\infty} h(l)h(l+k) \right] R_f(n-k). \quad (1.27)$$

Выражение (1.27), записанное с использованием оператора свертки, выглядит следующим образом:

$$R_g(n) = h(n) * h(-n) * R_f(n). \quad (1.28)$$

ВКФ входной и выходной последовательностей при $\mu_f = 0$ вычисляется в виде

$$\begin{aligned} R_{fg}(n) &= E \left\{ \left(f(n) - \mu_f \right) \left(g(n+k) - \mu_g \right) \right\} = \\ &= \sum_{k=-\infty}^{\infty} h(k) R_f(n-k) = h(n) * R_f(n). \end{aligned} \quad (1.29)$$

То есть искомая характеристика является сверткой импульсной характеристики ЛИС-системы и АКФ входного сигнала.

Энергетический спектр последовательности на выходе системы легко выводится из записанного выше соотношения (1.28). Действительно, с учетом свойств z -преобразования имеем

$$\Phi_g(z) = H(z)H(z^{-1})\Phi_f(z), \quad (1.30)$$

и далее, положив $z = e^{i\omega}$, получаем собственно энергетический спектр:

$$\Phi_g(e^{i\omega}) = H(e^{i\omega})H(e^{-i\omega})\Phi_f(e^{i\omega}). \quad (1.31)$$

Частотная характеристика обладает известной симметрией, поэтому выражение (1.31) может быть записано в более компактной форме:

$$\Phi_g(e^{i\omega}) = |H(e^{i\omega})|^2 \Phi_f(e^{i\omega}). \quad (1.32)$$

Получаем, что энергетический спектр последовательности на выходе ЛИС-системы равен энергетическому спектру входной последовательности, умноженному на квадрат модуля частотной характеристики системы.

Взаимный энергетический спектр входной и выходной последовательности вычисляется аналогично:

$$\Phi_{fg}(z) = H(z)\Phi_f(z), \quad (1.33)$$

и далее при $z = e^{i\omega}$:

$$\Phi_{fg}(e^{i\omega}) = H(e^{i\omega})\Phi_f(e^{i\omega}), \quad (1.34)$$

т.е. он равен произведению частотной характеристики системы и энергетического спектра входной последовательности.

1.4 Факторизация энергетического спектра

В развитие полученных результатов рассмотрим один важный методический прием, который часто используется при синтезе алгоритмов цифровой обработки сигналов.

Поставим следующую задачу: синтезировать физически реализуемую устойчивую ЛИС-систему, которая при поступлении на вход дискретного стационарного белого шума дает на выходе сигнал с заданной АКФ $R_g(n)$. Такую систему иногда называют “формирующим фильтром”. Для простоты изложения будем считать, что входной белый шум имеет единичную дисперсию, то есть его корреляционная функция

$$R_f(n) = \delta(n).$$

Нам известно выражение (1.30), связывающее энергетические спектры на входе и выходе ЛИС-системы. В данном случае $\Phi_f(z) = 1$, а энергетический спектр выходного сигнала $\Phi_g(z)$ – вычисляется по заданной последовательности $R_g(n)$. При этом вытекающее из (1.30) соотношение

$$\Phi_g(z) = H(z)H(z^{-1}) \quad (1.35)$$

можно рассматривать как уравнение относительно передаточной функции $H(z)$ искомого формирующего фильтра. Процедура нахождения $H(z)$ предполагает разложение $\Phi_g(z)$ на пару “симметричных” (в смысле (1.35)) множителей. Осуществление такого разложения будем называть факторизацией энергетического спектра.

Решение задачи факторизации не является единственным. Для того, чтобы оно имело практический смысл необходимо выполнить следующие два требования:

1) Найденная передаточная функция $H(z)$ должна соответствовать физически реализуемой ЛИС-системе конечного порядка, то есть допускать представление в дробно-рациональной форме, (в виде отношения полиномов по отрицательным степеням z).

2) Передаточная функция $H(z)$ должна соответствовать устойчивой ЛИС-системе, то есть иметь полюсы, лежащие внутри единичной окружности в комплексной z -плоскости.

Если энергетический спектр $\Phi_g(z)$ является дробно-рациональным, то среди решений задачи факторизации всегда найдется такое, которое удовлетворяет выдвинутым требованиям. Рассмотрим детально процедуру построения этого решения.

В силу четности АКФ $R_g(n)$ ее z -преобразование – энергетический спектр $\Phi_g(z)$ обладает свойством симметрии:

$$\Phi_g(z) = \Phi_g(z^{-1})$$

и, следовательно, если он является дробно-рациональным, может быть представлен в виде

$$\Phi_g(z) = \frac{B(z)}{A(z)}, \quad (1.36)$$

где

$$A(z) = \sum_{j=-M}^M a_j z^{-j}, \quad (1.37)$$

$$B(z) = \sum_{j=-N}^N b_j z^{-j} \quad (1.38)$$

полиномы из положительных и отрицательных степеней z с коэффициентами, удовлетворяющими условиям:

$$a_j = a_{-j}, \quad b_j = b_{-j}.$$

Рассмотрим сначала полином (1.37), стоящий в знаменателе дробно-рационального энергетического спектра (1.36). Уравнение

$$A(z) = 0$$

имеет $2M$ (то есть четное) число корней. Причем, благодаря симметрии коэффициентов, если комплексное число p – корень этого уравнения (полюс функции $\Phi_g(z)$), то и $1/p$ также является корнем (полюсом). Если $|p| < 1$, то $|1/p| > 1$, то есть половина корней будет лежать внутри единичной окружности комплексной

z -плоскости, а другая половина – вне единичной окружности. На самой единичной окружности корней нет, так как наличие таковых противоречило бы условиям сходимости рассматриваемого дробно-рационального z -преобразования при $|z|=1$). Обозначим через p_j ($1 \leq j \leq M$) корни, лежащие внутри единичной окружности. Несложно показать, что при этом степенной полином (1.37) может быть представлен через свои корни в виде:

$$A(z) = A_0 \prod_{j=1}^M (1 - p_j z^{-1}) \prod_{j=1}^M (1 - p_j z), \quad (1.39)$$

где A_0 – некоторая постоянная. Введем обозначение

$$A^+(z) = \sqrt{A_0} \prod_{j=1}^M (1 - p_j z^{-1}), \quad (1.40)$$

после которого выражение (1.39) записывается в форме

$$A(z) = A^+(z) A^+(z^{-1}), \quad (1.41)$$

то есть требуемая факторизация полинома $A(z)$ произведена.

Аналогичным образом осуществляется и факторизация полинома (1.38) из числителя дробно-рационального энергетического спектра:

$$B(z) = B^+(z) B^+(z^{-1}), \quad (1.42)$$

где

$$B^+(z) = \sqrt{B_0} \prod_{j=1}^M (1 - q_j z^{-1}) \quad (1.43)$$

– полином по отрицательным степеням z , B_0 – некоторая постоянная, q_j ($1 \leq j \leq N$) – корни $B^+(z)$.

Следует остановиться на особенностях выбора корней полинома (1.43). Во-первых, уравнение $B(z)=0$ может иметь решение, лежащее на единичной окружности комплексной z -плоскости, это всего лишь означает, что для некоторых частот ω энергетический спектр $\Phi_g(e^{i\omega})$ равен нулю. Во-вторых, к корням q_j ($1 \leq j \leq N$) нет необходимости предъявлять требование $|q_j| < 1$, поскольку, как мы увидим ниже, они будут определять положение нулей передаточной функции искомой ЛИС-системы, не влияющих на ее устойчивость. Основное условие формирования полинома (1.43): из всех $2N$ корней указанного уравнения в нём должно быть использовано по одному корню из каждой пары взаимобратных.

Полученные факторизованные представления (1.41) и (1.42) полиномов (1.37) и (1.38) позволяют произвести факторизацию и энергетического спектра (2.36) в целом:

$$\Phi_g(z) = \frac{B^+(z)B^+(z^{-1})}{A^+(z)A^+(z^{-1})} = \Phi_g^+(z)\Phi_g^+(z^{-1}), \quad (1.44)$$

где

$$\Phi_g^+(z) = \frac{B^+(z)}{A^+(z)} \quad (1.45)$$

дробно-рациональная функция от z^{-1} , не имеющая полюсов вне единичной окружности в z -плоскости. Из сравнения (1.44) с (1.35) видно, что в качестве искомой передаточной функции физически реализуемого и устойчивого формирующего фильтра можно принять

$$H(z) = \Phi_g^+(z)z^{-L}$$

при любом целом $L \geq 0$. Для простоты везде далее будем полагать $L = 0$, то есть брать

$$H(z) = \Phi_g^+(z) = \frac{B^+(z)}{A^+(z)}. \quad (1.46)$$

Пример 1.3. Определим передаточную функцию и построим разностное уравнение физически реализуемой и устойчивой ЛИС-системы, преобразующей белый шум с единичной дисперсией в стационарную случайную последовательность с АКФ

$$R_g(k) = \rho^{|k|} - a\rho^{|k-1|} - a\rho^{|k+1|}, \quad |\rho| < 1, |a| \leq 0.5.$$

Энергетический спектр выходной последовательности будет иметь вид:

$$\Phi_g(z) = \frac{(1-\rho^2)(1-az^{-1}-az)}{(1-\rho z^{-1})(1-\rho z)} = \frac{B(z)}{A(z)}.$$

Полином в знаменателе сразу записан в требуемой факторизованной форме:

$$A(z) = (1-\rho z^{-1})(1-\rho z) = A^+(z)A^+(z^{-1}),$$

где

$$A^+(z) = 1 - \rho z^{-1}.$$

Произведем факторизацию полинома в числителе, для чего решим уравнение

$$B(z) = (1-\rho^2)(1-az^{-1}-az) = 0$$

или

$$az^2 - z + a = 0.$$

Корни этого уравнения:

$$z_{1,2} = \frac{1}{2a} \left(1 \pm \sqrt{1 - 4a^2} \right).$$

Легко проверить, что они являются взаимобратными: $z_1 = \frac{1}{z_2}$. В зависимости от выбора одного из этих корней, используемого в качестве q_1 в (1.43), имеем два варианта факторизации $B(z)$:

$$B^+(z) = \sqrt{B_0} \left(1 - \frac{1 \pm \sqrt{1 - 4a^2}}{2a} z^{-1} \right),$$

где значения множителя

$$B_0 = (1 - \rho^2) \frac{1 \mp \sqrt{1 - 4a^2}}{2}$$

найлены подстановкой (1.43) в (1.42), раскрытием скобок и приравниванием коэффициента при любой из имеющихся степеней z к соответствующему коэффициенту в первоначальном представлении $B(z)$.

Итак, согласно (1.46), получаем две различные передаточные функции искомой ЛИС-системы:

$$H(z) = \sqrt{(1 - \rho^2) \frac{1 \mp \sqrt{1 - 4a^2}}{2}} \frac{\left(1 - \frac{1 \pm \sqrt{1 - 4a^2}}{2a} z^{-1} \right)}{1 - \rho z^{-1}},$$

по которым легко строятся два варианта описывающих систему разностных уравнений:

$$g(n) = \rho g(n-1) + \sqrt{(1-\rho^2)} \frac{1 \mp \sqrt{1-4a^2}}{2} \left[f(n) - \frac{1 \pm \sqrt{1-4a^2}}{2a} f(n-1) \right].$$

Процедуру факторизации, очевидно, можно использовать и для решения более общей задачи, чем та, которая была поставлена в начале данного параграфа, а именно для синтеза физически реализуемой устойчивой ЛИС-системы, преобразующей стационарную случайную последовательность с одной АКФ $R_f(n)$ в последовательность с другой АКФ $R_g(n)$. Действительно, непосредственно из (1.30) следует

$$\frac{\Phi_g(z)}{\Phi_f(z)} = H(z)H(z^{-1}),$$

дробно-рациональные энергетические спектры входного и выходного сигналов могут быть факторизованы:

$$\Phi_f(z) = \Phi_f^+(z) \Phi_f^+(z^{-1}),$$

$$\Phi_g(z) = \Phi_g^+(z) \Phi_g^+(z^{-1}),$$

где

$$\Phi_f^+(z) = \frac{B_f^+(z)}{A_f^+(z)}, \quad \Phi_g^+(z) = \frac{B_g^+(z)}{A_g^+(z)},$$

$A_f^+(z)$, $B_f^+(z)$, $A_g^+(z)$, $B_g^+(z)$ – полиномы, определяемые в процессе факторизации, и, следовательно, в качестве передаточной функции ЛИС-системы можно принять:

$$H(z) = \frac{\Phi_g^+(z)}{\Phi_f^+(z)} = \frac{A_f^+(z)B_g^+(z)}{A_g^+(z)B_f^+(z)}. \quad (1.47)$$

Но здесь нужно более строго, чем раньше, подходить к выбору корней при факторизации числителя энергетического спектра входного сигнала – $\Phi_f(z)$, то есть при конструировании полинома $B_f^+(z)$: в соответствии с (1.47) корни этого полинома оказываются полюсами передаточной функции и для того, чтобы система была устойчивой, они должны обязательно выбираться внутри единичной окружности z -плоскости. Задача не будет иметь решения (система не получится устойчивой), если у $B_f^+(z)$ будут иметься корни, лежащие на единичной окружности, и эти корни не будут скомпенсированы соответствующими корнями $B_g^+(z)$.

2 ОДНОМЕРНАЯ ЛИНЕЙНАЯ ФИЛЬТРАЦИЯ И ВОССТАНОВЛЕНИЕ СИГНАЛОВ

2.1 Восстановление дискретного сигнала ЛИС-системой

Пусть имеется полезный сигнал – последовательность $f(n)$. Однако непосредственному наблюдению (измерению) он недоступен. В нашем распоряжении имеется лишь сигнал $g(n)$ – результат прохождения сигнала через некоторую «искажающую» систему, дополнительно искаженный шумом $v(n)$ (см. рисунок 2.1).



Рис. 2.1. Модель наблюдения полезного сигнала

Требуется восстановить полезный сигнал по наблюдаемому. Для этого необходимо синтезировать такую восстанавливающую систему (фильтр), чтобы при подаче на ее вход наблюдаемого сигнала на выходе получалась бы оценка $\hat{f}(n)$ полезного сигнала (см. рисунок 2.2).

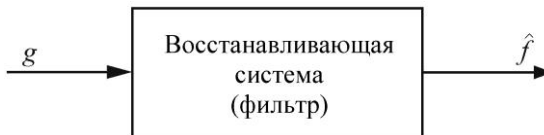


Рис. 2.2. Схема восстанавливающей системы

Далее мы сузим класс рассматриваемых сигналов и систем.

Во-первых, в большинстве практически важных случаев искажения сигнала удастся описать моделью ЛИС-системы, рассмотрением которой мы и ограничимся. Будем считать, что известна ее импульсная характеристика $h(n)$. Тогда наблюдаемая последовательность запишется в виде:

$$g(n) = f(n) * h(n) + v(n). \quad (2.1)$$

Соотношение (2.1) задает так называемую *линейную модель наблюдения в дискретном времени*.

Во-вторых, восстанавливать сигнал будем также при помощи ЛИС-системы:

$$\hat{f}(n) = g(n) * h_{восст}(n), \quad (2.2)$$

где $h_{восст}(n)$ – импульсная характеристика восстанавливающей ЛИС-системы.

В-третьих, и полезный сигнал $f(n)$, и шум $v(n)$ будем считать стационарными случайными последовательностями, статистические характеристики которых известны.

Заметим, что, поскольку все преобразуемые последовательности случайны, то и ошибка восстановления:

$$\varepsilon(n) = \hat{f}(n) - f(n) \quad (2.3)$$

в каждый момент времени тоже случайна. Мы будем строить такой восстанавливающий фильтр, который обеспечивает минимум ошибки в среднеквадратичном смысле, то есть минимизирует ее дисперсию:

$$\bar{\varepsilon}^2 = E \left\{ \varepsilon^2(n) \right\} = E \left\{ \left[\hat{f}(n) - f(n) \right]^2 \right\} \rightarrow \min. \quad (2.4)$$

Из всего сказанного наиболее существенным является ограничение, заключающееся в требовании линейности восстанавливающей системы. Однако для нелинейных систем получить конкретные результаты их синтеза гораздо сложнее. Кроме того, из теории информации известно, что для важного класса сигналов – гауссовских – оптимальное (наилучшее) в среднеквадратичном смысле восстановление обеспечивается именно линейной системой.

ЛИС-система, реализующая преобразование (2.2) и обеспечивающая при этом выполнение условия (2.4), называется «*оптимальным линейным восстанавливающим фильтром*». А ее применение реализует процедуру оптимального линейного восстановления.

Очень часто, однако, на импульсную характеристику восстанавливающей ЛИС-системы налагаются дополнительные ограничения, связанные с удобством реализации. Например, требуется, чтобы она была КИХ-системой или физически реализуемой БИХ-системой (см. [1]). В таких ситуациях ошибка восстановления несколько возрастёт, то есть мы получим *квазиоптимальные* процедуры восстановления.

Мы объединим рассмотрение оптимального и квазиоптимального восстановления следующим образом: будем считать, что импульсная характеристика восстанавливающей ЛИС-системы отлична от нуля для значений аргумента из некоторого множества D (интервала наблюдения):

$$h_{восст}(n) = 0 \text{ при } n \notin D. \quad (2.5)$$

Определим при этом ограничение параметры системы, минимизирующие ошибку восстановления. С учетом сказанного выше конкретизируется формула (2.2):

$$\hat{f}(n) = \sum_{k=-\infty}^{\infty} h_{восст}(k) g(n-k) = \sum_{k \in D} h_{восст}(k) g(n-k) \quad (2.6)$$

и условие минимизации ошибки (2.4):

$$\begin{aligned}\bar{\varepsilon}^2 &= E \left\{ \left[\tilde{f}(n) - f(n) \right]^2 \right\} \\ &= E \left\{ \left[\sum_{k \in D} h_{\text{восст}}(k) g(n-k) - f(n) \right]^2 \right\} \rightarrow \min.\end{aligned}\quad (2.7)$$

Минимизация ошибки осуществляется путем варьирования ненулевых отсчетов импульсной характеристики восстанавливающей системы. В точке минимума обеспечивается равенство нулю всех частных производных:

$$\frac{\partial \varepsilon^2}{\partial h_{\text{восст}}(m)} = 0, \quad m \in D. \quad (2.8)$$

Подставив (2.7) в (2.8), получаем:

$$\begin{aligned}&\frac{\partial \varepsilon^2}{\partial h_{\text{восст}}(m)} = \\ &= E \left\{ \left[\sum_{k \in D} h_{\text{восст}}(k) g(n-k) - f(n) \right] g(n-m) \right\} = 0, \\ &\quad m \in D.\end{aligned}\quad (2.9)$$

Из последнего выражения следуют два важных соотношения.

Во-первых, это выражение можно записать в виде:

$$E \left\{ \varepsilon(n) g(n-m) \right\} = 0, \quad m \in D, \quad (2.10)$$

то есть ВКФ $R_{\varepsilon g}(m) = 0$, ошибка оптимального восстановления некоррелирована с наблюдаемым сигналом. Это утверждение известно в теории оценивания как «*лемма об ортогональном проектировании*», которая будет нам полезна в дальнейшем.

Во-вторых, перенеся в (2.9) вычитаемое в правую часть после применения оператора математического ожидания, получим:

$$\sum_{k \in D} h_{\text{восст}}(k) R_g(m-k) = R_{fg}(-m) \quad (2.11)$$

– уравнение Винера-Хопфа для дискретных систем.

Таким образом, импульсная характеристика оптимального (или квазиоптимального) линейного восстанавливающего фильтра определяется из системы уравнений, состоящей из уравнения Винера-Хопфа и ограничений, налагаемых на импульсную характеристику:

$$\begin{cases} \sum_{k \in D} h_{\text{восст}}(k) R_g(m-k) = R_{fg}(-m), & m \in D, \\ h_{\text{восст}}(m) = 0, & m \notin D. \end{cases} \quad (2.12)$$

Различный вид области D приводит к существенно различным методам решения системы (2.12), которые будут описаны ниже. Но сначала определим минимизируемую ошибку восстановления сигнала, продолжив преобразования, входящие в (2.7), с учетом (2.3) и (2.10):

$$\begin{aligned} \bar{\varepsilon}^2 &= E \left\{ \varepsilon(n) \left[\sum_{k \in D} h_{\text{восст}}(k) g(n-k) - f(n) \right] \right\} = \\ &= E \left\{ \varepsilon(n) \sum_{k \in D} h_{\text{восст}}(k) g(n-k) \right\} - E \{ \varepsilon(n) f(n) \} = \\ &= \sum_{k \in D} h_{\text{восст}}(k) E \{ \varepsilon(n) g(n-k) \} - E \{ \varepsilon(n) f(n) \} = \\ &= -E \{ \varepsilon(n) f(n) \} = -E \left\{ \left[\sum_{k \in D} h_{\text{восст}}(k) g(n-k) - f(n) \right] f(n) \right\} \\ &= E \{ f^2(n) \} - \sum_{k \in D} h_{\text{восст}}(k) E \{ f(n) g(n-k) \} = \\ &= \sigma_f^2 - \sum_{k \in D} h_{\text{восст}}(k) R_{fg}(-k). \end{aligned} \quad (2.13)$$

Рассмотрим важный частный случай, когда имеет место упрощенная модель наблюдения с белым шумом, независимым от сигнала:

$$g(n) = f(n) + v(n), \quad R_v(k) = \sigma_v^2 \delta(k) \quad (2.14)$$

и нулевой отсчет импульсной характеристики $h_{\text{восстм}}(0)$ не равен нулю, ($\{0\} \in D$). В этом случае

$$\begin{aligned} R_g(k) &= R_f(k) + R_v(k) = R_f(k) + \sigma_v^2 \delta(k), \\ R_{fg}(k) &= E\{f(n)g(n+k)\} = E\{f(n)[f(n+k)] + v(n+k)\} = \\ &= E\{f(n)f(n+k)\} + E\{f(n)v(n+k)\} = R_f(k) = R_f(-k). \end{aligned}$$

Тогда из уравнения Винера-Хопфа (2.11) получаем:

$$\begin{aligned} \sum_{k \in D} h_{\text{восстм}}(k) [R_f(m-k) + \sigma_v^2 \delta(m-k)] &= R_f(m), \\ \sum_{k \in D} h_{\text{восстм}}(k) R_f(m-k) + \sigma_v^2 h_{\text{восстм}}(m) &= R_f(m) \end{aligned}$$

и при $m = 0$:

$$\sum_{k \in D} h_{\text{восстм}}(k) R_f(-k) + \sigma_v^2 h_{\text{восстм}}(0) = \sigma_f^2. \quad (2.15)$$

При этих же условиях выражение (2.13) для ошибки восстановления приобретает вид:

$$\varepsilon^2 = \sigma_f^2 - \sum_{k \in D} h_{\text{восстм}}(k) R_f(-k) = \sigma_v^2 h_{\text{восстм}}(0),$$

а после подстановки в него выражения (2.15):

$$\varepsilon^2 = \sigma_v^2 h_{\text{восстм}}(0). \quad (2.16)$$

Это простое соотношение нам будет полезно в дальнейшем.

2.2 Оптимальное линейное восстановление

Пусть на отсчеты импульсной характеристики восстанавливающей ЛИС-системы не наложено никаких ограничений, то есть она может быть отлична от нуля в любой точке. Это значит, что в оценке полезного сигнала будут учтены все наблюдаемые отсчеты (как “прошлые”, так и “будущие”). При этом восстановление, очевидно, будет наилучшим (оптимальным).

Так как ограничений на $h_{восст}(n)$ в данном случае нет, то из введенной в предыдущем параграфе системы уравнений (2.12) остается только уравнение Винера-Хопфа, записываемое в виде:

$$\sum_{k=-\infty}^{\infty} h_{восст}(k) R_g(m-k) = R_{fg}(-m). \quad (2.17)$$

Выражение (2.17) можно интерпретировать как свертку последовательностей, поэтому, переходя к их z -преобразованиям, получаем:

$$\begin{aligned} H_{восст}(z) \Phi_g(z) &= \Phi_{fg}(z^{-1}), \\ H_{восст}(z) &= \frac{\Phi_{fg}(z^{-1})}{\Phi_g(z)}. \end{aligned} \quad (2.18)$$

Формула (2.18) задает передаточную функцию искомого оптимального фильтра. Его импульсная характеристика может быть определена отсюда обычным путем через обратное z -преобразование.

Определим, какую минимальную ошибку восстановления обеспечивает оптимальный фильтр. Для этого можно было бы, конечно, воспользоваться формулой (2.13), полученной в предыдущем параграфе, но в данном случае удобнее и полезнее для анали-

за сделать иначе. Определим сначала АКФ и энергетический спектр ошибки восстановления:

$$\begin{aligned}
 R_{\varepsilon}(m) &= E\{\varepsilon(n)\varepsilon(n+m)\} = \\
 &= E\left\{\left[\sum_{k=-\infty}^{\infty} h_{восст}(k)g(n-k) - f(n)\right] \times \right. \\
 &\quad \left. \times \left[\sum_{l=-\infty}^{\infty} h_{восст}(l)g(n+m-l) - f(n+m)\right]\right\} = \\
 &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h_{восст}(k)h_{восст}(l)R_g(m-l+k) - \\
 &\quad - \sum_{k=-\infty}^{\infty} h_{восст}(k)R_{fg}(-k-m) - \\
 &\quad - \sum_{l=-\infty}^{\infty} h_{восст}(l)R_{fg}(m-l) + R_f(m).
 \end{aligned}$$

Запишем это же выражение в сокращенной форме, используя оператор свертки:

$$\begin{aligned}
 R_{\varepsilon}(m) &= h_{восст}(m) * h_{восст}(-m) * R_g(m) - \\
 &\quad - h_{восст}(-m) * R_{fg}(-m) - \\
 &\quad - h_{восст}(m) * R_{fg}(m) + R_f(m)
 \end{aligned} \tag{2.19}$$

и перейдем к z -преобразованиям последовательностей, входящих в (2.19):

$$\begin{aligned}
 \Phi_{\varepsilon}(z) &= H_{восст}(z)H_{восст}(z^{-1})\Phi_g(z) - \\
 &\quad - H_{восст}(z^{-1})\Phi_{fg}(z^{-1}) - \\
 &\quad - H_{восст}(z)\Phi_{fg}(z) + \Phi_f(z).
 \end{aligned} \tag{2.20}$$

Выражения (2.19) и (2.20) справедливы для любой восстанавливающей системы, а не только для оптимальной, поскольку при

их получении мы не налагали ограничений на $h_{восст}(n)$. Для оптимального фильтра учтем соотношение (2.18) и получим:

$$\Phi_{\varepsilon}(z) = \Phi_f(z) - \frac{\Phi_{fg}(z^{-1})}{\Phi_g(z)} \Phi_{fg}(z)$$

или, что удобнее,

$$\Phi_{\varepsilon}(z) = \Phi_f(z) - H_{восст}(z) \Phi_{fg}(z). \quad (2.21)$$

Из последней формулы искомую среднеквадратичную ошибку можно вычислить двумя путями:

1. Перейти от z -преобразования (2.21) к самой последовательности (корреляционной функции ошибки):

$$R_{\varepsilon}(m) = R_f(m) - \sum_{k=-\infty}^{\infty} h_{восст}(k) R_{fg}(m-k)$$

и далее при $m=0$ получить

$$\varepsilon^2 = \sigma_f^2 - \sum_{k=-\infty}^{\infty} h_{восст}(k) R_{fg}(-k),$$

(последняя формула является частным случаем формулы (2.13)).

2. Перейти от z -преобразования к энергетическому спектру ошибки восстановления и вычислить ошибку по формуле:

$$\varepsilon^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{\varepsilon}(e^{i\omega}) d\omega.$$

В общей постановке задачи решение на этом завершается. Более продвинутой результат можно получить, введя дополнительные упрощения.

Рассмотрим частную, но широко распространенную ситуацию восстановления сигнала при линейной модели наблюдения (2.1),

когда полезный сигнал и шум статистически независимы. Оптимальный линейный восстанавливающий фильтр для этого случая называется *фильтром Винера-Колмогорова*. Определим его передаточную функцию. Для начала подсчитаем корреляционные функции, входящие в (2.17). АКФ наблюдаемого сигнала:

$$\begin{aligned}
 R_g(m) &= E\{g(n)g(n+m)\} = \\
 &= E\left\{\left[\sum_{k=-\infty}^{\infty} h(k)f(n-k)+v(n)\right] \times \right. \\
 &\quad \left. \times \left[\sum_{l=-\infty}^{\infty} h(l)f(n+m-l)+v(n+m)\right]\right\} = \\
 &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k)h(l)E\{f(n-k)f(n+m-l)\} + \\
 &\quad + \sum_{k=-\infty}^{\infty} h(k)E\{f(n-k)v(n+m)\} + \\
 &\quad + \sum_{l=-\infty}^{\infty} h(l)E\{f(n+m-l)v(n)\} + E\{v(n)v(n+m)\} = \\
 &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k)h(l)R_f(n+m-l) + R_v(m) = \\
 &= h(m) * h(-m) * R_f(m) + R_v(m).
 \end{aligned}$$

Соответственно, для z -преобразований записанных последовательностей:

$$\Phi_g(z) = H(z)H(z^{-1})\Phi_f(z) + \Phi_v(z). \quad (2.22)$$

ВКФ полезного и наблюдаемого сигнала получается в результате аналогичных, но более простых преобразований:

$$\begin{aligned}
 R_{fg}(m) &= E\{f(n)g(n+m)\} = \\
 &= \sum_{k=-\infty}^{\infty} h(k)R_f(m-k) = h(m) * R_f(m) ,
 \end{aligned}$$

то есть

$$\Phi_{fg}(z) = H(z)\Phi_f(z). \quad (2.23)$$

Подставив (2.22) и (2.23) в (2.18), получаем передаточную функцию фильтра Винера-Колмогорова:

$$H_{восст}(z) = \frac{H(z^{-1})\Phi_f(z)}{H(z)H(z^{-1})\Phi_f(z) + \Phi_v(z)} \quad (2.24)$$

(в этой формуле дополнительно учтено, что $\Phi_f(z^{-1}) = \Phi_f(z)$).

Фильтр Винера-Колмогорова обеспечивает минимальную среднеквадратичную ошибку восстановления сигнала при линейной модели наблюдения и отсутствии корреляции между полезным сигналом и шумом. Энергетический спектр этой ошибки можно найти подстановкой (2.23) и (2.24) в (2.21):

$$\Phi_\varepsilon(z) = \frac{\Phi_v(z)\Phi_f(z)}{H(z)H(z^{-1})\Phi_f(z) + \Phi_v(z)}, \quad (2.25)$$

а сама ошибка определяется отсюда известными путями, описанными выше.

Рассмотрим некоторые частные случаи применения фильтра Винера-Колмогорова.

1. Пусть имеется упрощенная модель наблюдения без “линейных” искажений:

$$g(n) = f(n) + v(n). \quad (2.26)$$

Здесь $h(n) = \delta(n)$, $H(z) = 1$ и поэтому из (2.24), (2.25) получаем:

$$H_{\text{восст}}(z) = \frac{\Phi_f(z)}{\Phi_f(z) + \Phi_v(z)}, \quad (2.27)$$

$$\Phi_\varepsilon(z) = \frac{\Phi_v(z)\Phi_f(z)}{\Phi_f(z) + \Phi_v(z)} = \Phi_v(z)H_{\text{восст}}(z). \quad (2.28)$$

Нетрудно заметить, что в данном случае, поскольку

$$\Phi_f(z) = \Phi_f(z^{-1}); \quad \Phi_v(z) = \Phi_v(z^{-1}),$$

то и

$$H_{\text{восст}}(z) = H_{\text{восст}}(z^{-1}),$$

а это означает, что

$$h_{\text{восст}}(n) = h_{\text{восст}}(-n)$$

– импульсная характеристика фильтра является четной последовательностью. Такой фильтр является физически нереализуемым за исключением единственного вырожденного случая, рассматриваемом ниже.

2. Пусть кроме того шум – белый, то есть

$$R_v(m) = \sigma_v^2 \delta(m); \quad \Phi_v(z) = \sigma_v^2.$$

Тогда из (2.27), (2.28) имеем:

$$H_{\text{восст}}(z) = \frac{\Phi_f(z)}{\Phi_f(z) + \sigma_v^2}, \quad (2.29)$$

$$\Phi_\varepsilon(z) = \frac{\sigma_v^2 \Phi_f(z)}{\Phi_f(z) + \sigma_v^2}. \quad (2.30)$$

3. Пусть, наконец, и полезный сигнал также является белым шумом (этот случай, как мы увидим ниже, имеет определенный практический смысл). Теперь

$$R_f(m) = \sigma_f^2 \delta(m) ; \quad \Phi_f(z) = \sigma_f^2 . \quad (2.31)$$

и, следовательно,

$$H_{восст}(z) = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} ; \quad \Phi_\varepsilon(z) = \frac{\sigma_v^2 \sigma_f^2}{\sigma_f^2 + \sigma_v^2} . \quad (2.32)$$

От (2.32) можно легко перейти во временную область:

$$h_{восст}(n) = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} \delta(n) ,$$

и далее записать:

$$\hat{f}(n) = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} g(n) ,$$

то есть фильтрация заключается в простом умножении наблюдаемого сигнала на коэффициент (это так называемая “точечная” оценка сигнала).

Ошибки восстановления в соответствии с (2.32):

$$R_\varepsilon(m) = \frac{\sigma_v^2 \sigma_f^2}{\sigma_f^2 + \sigma_v^2} \delta(m) ; \quad \bar{\varepsilon}^2 = R_\varepsilon(0) = \frac{\sigma_v^2 \sigma_f^2}{\sigma_f^2 + \sigma_v^2} ,$$

то есть дисперсия входного шума σ_v^2 здесь умножается на коэффициент $\frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} < 1$ – шум убывает в максимально достижимой степени.

4. Еще один частный случай – отсутствие шума. При этом из (2.24) получаем:

$$H_{\text{восст}}(z) = \frac{H(z^{-1})\Phi_f(z)}{H(z)H(z^{-1})\Phi_f(z)} = \frac{1}{H(z)}$$

– так называемый *обратный (инверсный) фильтр*. В идеале такой фильтр обеспечивает абсолютно точное восстановление сигнала. Однако в большинстве практически интересных случаев он оказывается неустойчивым: бесконечно малым отклонением входного сигнала обратного фильтра могут соответствовать бесконечно большие отклонения выходного сигнала, то есть *задача восстановления относится к числу некорректных*. Для получения устойчивого фильтра используются различные *методы регуляризации*.

Пример 2.1. Пусть модель наблюдения сигнала имеет вид (2.26), полезный сигнал имеет экспоненциальную АКФ:

$$R_f(m) = \sigma_f^2 \rho^{|m|},$$

где ρ – коэффициент корреляции между соседними отсчетами, и наблюдается на фоне белого шума:

$$R_v(m) = \sigma_v^2 \delta(m).$$

Определим передаточную функцию фильтра Винера. В данном случае:

$$\Phi_f(z) = \frac{\sigma_f^2(1-\rho^2)}{(1-\rho z^{-1})(1-\rho z)}, \quad \Phi_v(z) = \sigma_v^2,$$

и, подставив эти величины в (2.27), после преобразований получаем:

$$H_{\text{восст}}(z) = \frac{d^2(1-\rho^2)}{d^2(1-\rho^2) + (1+\rho^2) - \rho(z+z^{-1})}, \quad (2.33)$$

где обозначено $d^2 = \frac{\sigma_f^2}{\sigma_v^2}$ – отношение сигнал/шум по мощности.

Поскольку фильтр должен быть устойчивым, область сходимости данного z -преобразования должна включать в себя единичную окружность.

Из (2.33), основываясь на свойствах z -преобразования, можно определить импульсную характеристику фильтра Винера-Колмогорова:

$$h_{восст}(n) = A\alpha^{|n|}, \quad (2.34)$$

где

$$A = \frac{1}{\sqrt{\left(1 + \frac{1}{d^2}\right)^2 + \frac{4\rho^2}{d^2(1-\rho^2)}}},$$

$$\alpha = \frac{1}{2} \left\{ \left(\frac{1}{\rho} + \rho \right) + d^2 \left(\frac{1}{\rho} - \rho \right) \left[\sqrt{\left(1 + \frac{1}{d^2}\right)^2 + \frac{4\rho^2}{d^2(1-\rho^2)}} \right] \right\}.$$

Можно показать, что всегда $A > 0$; $|\alpha| < 1$.

Фильтр с импульсной характеристикой вида (2.34), очевидно является физически нереализуемым. Поэтому вопрос его практического использования пока остается открытым. Ответ на него мы дадим позже. А пока определим ошибку восстановления. В нашем случае линейных искажений сигнала нет и шум белый, поэтому сразу можно воспользоваться формулой (2.16):

$$\bar{\varepsilon}^2 = \sigma_v^2 h_{восст}(0) = \frac{\sigma_v^2}{\sqrt{\left(1 + \frac{1}{d^2}\right)^2 + \frac{4\rho^2}{d^2(1-\rho^2)}}}. \quad (2.35)$$

Выражение (2.35) имеет смысл проанализировать. При $\rho \rightarrow 1$, то есть при увеличении коррелированности полезного сигнала, $\varepsilon^2 \rightarrow 0$ и возможность фильтрации шума возрастает. При увеличении отношения сигнал/шум ($d^2 \rightarrow \infty$) отношение $\frac{\varepsilon^2}{\sigma_v^2} \rightarrow 1$, и относительная эффективность фильтрации (коэффициент подавления шума) стремится к единице. Иными словами, слабый шум фильтруется плохо. Иллюстрация к сказанному дана на рисунке 2.3.

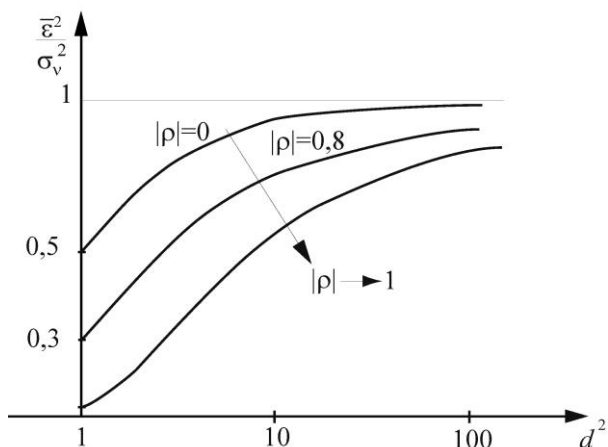


Рис. 2.3. Зависимость качества восстановления от параметров искажения

2.3 Реализация оптимального фильтра обработкой «в прямом и обратном времени»

Оптимальный линейный восстанавливающий фильтр, как правило, не отвечает требованию физической реализуемости. Поэтому оценка сигнала (2.2) не может быть вычислена напрямую. Для того, чтобы практически воспользоваться процедурой оптимального

восстановления, есть два основных способа. В данном параграфе мы рассмотрим один из них, заключающийся в *обработке сигнала «в прямом и обратном времени»*.

Этот способ обработки применяется в тех случаях, когда есть возможность сразу ввести в компьютер достаточно длинную реализацию сигнала. Когда отсчеты последовательности записаны в память компьютера, понятия “прошлого” и “будущего” становятся условными: по сигналу (то есть по массиву отсчетов) можно двигаться как в направлении возрастания аргумента (индекса), то есть в *«прямом времени»*, так и в направлении убывания – *«в обратном времени»*. Этот факт и позволяет реализовать оптимальный фильтр.

Ниже будем считать, что характеристики обрабатываемых сигналов таковы, что передаточная функция оптимального фильтра $H_{восст}(z)$ является дробно-рациональной. Она соответствует *устойчивой, но физически нереализуемой системе*, то есть взаимное расположение полюсов и области сходимости на z -плоскости имеет примерно такой вид, как изображенный на рисунке 2.4.

Областью сходимости $H_{восст}(z)$ является кольцо, включающее единичную окружность:

$$R_- < |z| < R_+ \text{ при } R_- < 1; R_+ > 1 .$$

Такую дробно-рациональную передаточную функцию можно записать через нули и полюсы (см. также формулу (4.10) в [1]):

$$H_{восст}(z) = A \frac{\prod_{j=1}^N (1 - q_j z^{-1})}{\prod_{j=1}^M (1 - p_j z^{-1})} \cdot z^k , \quad (2.36)$$

где A, k – некоторые константы (k – целое). Часть полюсов в (2.36) имеет модуль меньше единицы, а часть – больше единицы.

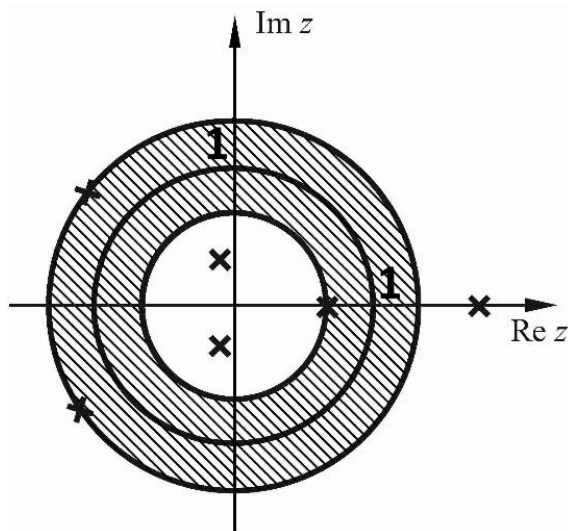


Рис. 2.4. Расположение полюсов устойчивой физически нереализуемой системе

Представим передаточную функцию в следующей форме:

$$H_{\text{восст}}(z) = H_{\text{восст}}^+(z) H_{\text{восст}}^-(z), \quad (2.37)$$

где к сомножителю $H_{\text{восст}}^+(z)$ отнесем часть знаменателя с полюсами, лежащими внутри единичной окружности, а к $H_{\text{восст}}^-(z)$ – с полюсами вне единичной окружности. Распределение нулей и коэффициента A , в принципе, произвольно. Очевидно, что здесь мы снова решаем задачу факторизации (см. п. 1.3), но уже в более общей, “несимметричной” постановке.

Составляющая $H_{\text{восст}}^+(z)$ будет иметь область сходимости

$$|z| > R_- \quad (R_- < 1),$$

то есть соответствовать передаточной функции некоторой устойчивой системы. Эта система физически реализуема, так как ее импульсная характеристика $h_{восст}^+(n)$, соответствующая передаточной функции $H_{восст}^+(z)$, является правосторонней последовательностью.

Аналогично, сомножитель $H_{восст}^-(z)$ в (2.37) имеет область сходимости

$$|z| < R_+ \quad (R_+ > 1),$$

и соответствует передаточной функции устойчивой системы, реализуемой в обратном времени (ее импульсная характеристика $h_{восст}^-(n)$ будет левосторонней последовательностью).

Произведение передаточных функций соответствует каскадному (последовательному) соединению ЛИС-систем. То есть мы имеем здесь “двухпроходную” процедуру восстановления, заключающуюся в последовательной обработке сигнала в прямом, и затем в обратном времени.

С другой стороны, можно представить передаточную функцию $H_{восст}(z)$ в виде суммы, используя разложение (2.36) на простые дроби (см. также формулу (4.37) в [1]):

$$H_{восст}(z) = \sum_{j=1}^M \frac{C_j}{1 - p_j z^{-1}} \quad (2.38)$$

(выражение (2.38) записано для случая правильной дроби и простых полюсов).

В данном случае получаем

$$H_{восст}(z) = H_{восст}^+(z) + H_{восст}^-(z). \quad (2.39)$$

где слагаемые формируются по тому же принципу, что и сомножители в (2.37). Формула (2.39) задает двухпроходную процедуру параллельной обработки сигнала.

Пример 2.2. В предыдущем параграфе мы получали, что для восстановления сигнала с экспоненциальной АКФ из его смеси с независимым белым шумом импульсная характеристика оптимального (винеровского) фильтра имеет вид:

$$h_{восст}(n) = A\alpha^{|n|},$$

где $A > 0$, $|\alpha| < 1$ – величины, рассчитываемые через характеристики сигнала и шума.

Передаточная функция этого фильтра:

$$H_{восст}(z) = \frac{A(1-\alpha^2)}{(1-\alpha z^{-1})(1-\alpha z)}$$

с полюсами $p_1 = \alpha$, $p_2 = \frac{1}{\alpha}$. Построим двухпроходный последовательный алгоритм обработки. В данном случае передаточная функция легко факторизуется к виду (2.37), где, например,

$$H_{восст}^+(z) = \frac{1}{1-\alpha z^{-1}}; \quad H_{восст}^-(z) = \frac{A(1-\alpha^2)}{1-\alpha z}.$$

По этим передаточным функциям строятся разностные уравнения. На первом шаге обработки (в прямом времени) из искаженного сигнала $g(n)$ получаем промежуточную последовательность $w(n)$:

$$w(n) = \alpha w(n-1) + g(n).$$

На втором шаге обработки (в обратном времени) получаем искомую оценку сигнала:

$$\hat{f}(n) = \alpha \hat{f}(n+1) + w(n)A(1-\alpha^2).$$

Можно построить и двухпроходный параллельный алгоритм. Для этого, вообще говоря, нужно разложить передаточную функцию $H_{восст}(z)$ на простые дроби. Но в данном конкретном случае можно поступить проще и сразу представить импульсную характеристику фильтра в следующем виде:

$$h_{восст}(n) = A\alpha^{|n|} = A\left[\alpha^n u(n) + \alpha^{-n} u(-n) - \delta(n)\right],$$

то есть, например,

$$H_{восст}(z) = A\left[\frac{1}{1-\alpha z^{-1}} + \frac{1}{1-\alpha z} - 1\right] = H_{восст}^+(z) + H_{восст}^-(z),$$

где

$$H_{восст}^+(z) = A\left[\frac{1}{1-\alpha z^{-1}} - 1\right] = \frac{A\alpha z^{-1}}{1-\alpha z^{-1}},$$

$$H_{восст}^-(z) = \frac{A}{1-\alpha z}.$$

В соответствии с полученными соотношениями, при обработке в прямом времени формируется последовательность $\hat{f}^+(n)$:

$$\hat{f}^+(n) = \alpha \hat{f}^+(n-1) + A\alpha g(n-1),$$

а при обработке в обратном времени – $\hat{f}^-(n)$:

$$\hat{f}^-(n) = \alpha \hat{f}^-(n+1) + g(n).$$

Далее для получения результата восстановления эти последовательности суммируются:

$$\hat{f}(n) = \hat{f}^+(n) + \hat{f}^-(n).$$

2.4 Реализация оптимального фильтра при помощи ДПФ

Оптимальный линейный фильтр физически реализуем и при этом чрезвычайно прост в ситуации восстановления белого шума на фоне белого шума, сводящегося, как мы видели, к точечной оценке (2.32). В общем случае сигнал не является белым шумом, в нём наблюдается статистическая связь между отсчетами, поэтому при решении задачи восстановления мы приходим к уравнению Винера-Хопфа. Однако есть и другая возможность построения процедуры оптимального восстановления. Можно произвести над сигналом некоторое обратимое преобразование, которое произвело бы декорреляцию сигнала. К декоррелированному сигналу можно применить процедуру точечной оценки, которая для такой ситуации будет оптимальной. Затем после обратного преобразования получим искомую оценку сигнала.

Требуемым декоррелирующим свойством при определенных условиях обладает дискретное преобразование Фурье (ДПФ), детально рассмотренное в главе 3 учебного пособия [1] (см. там формулы (3.5) и (3.9)). Рассмотрим более подробно процедуру оптимального восстановления в спектральной области на примере, когда имеется модель наблюдения без динамических искажений, заданная соотношением (2.26).

Поскольку ДПФ предполагает работу с последовательностями конечной длины, наблюдаемый сигнал разбивается на отрезки длиной по N отсчетов. Рассмотрим один из таких отрезков при $0 \leq n \leq N-1$. После применения ДПФ к (2.26) получаем уравнение наблюдения для дискретных спектров:

$$G(m) = F(m) + V(m), \quad 0 \leq m \leq N-1. \quad (2.40)$$

Поскольку последовательности в исходной модели наблюдения считаются случайными, их ДПФ тоже являются случайными

последовательностями. И для восстановления сигнала нам нужно знать их статистические характеристики.

Далее все количественные соотношения и формулы получим для нашего сквозного примера из п.2.2 и п.2.3: будем считать, что экспоненциально коррелированный сигнал наблюдается на фоне белого шума, то есть

$$R_f(k) = \sigma_f^2 \rho^{|k|}, \quad (2.41)$$

$$R_v(k) = \sigma_v^2 \delta(k). \quad (2.42)$$

Найдем АКФ преобразованного по ДПФ полезного сигнала. По определению, для нестационарной комплексной случайной последовательности ($0 \leq k, l \leq N-1$) ее АКФ задается выражением

$$B_F(k, l) = E \left\{ F(k) F^*(l) \right\}, \quad (2.43)$$

где * – знак комплексного сопряжения. Подставив в (2.43) сначала выражение для прямого ДПФ, а затем (2.41), получаем:

$$\begin{aligned} R_F(k, l) &= E \left\{ \sum_{p=0}^{N-1} f(p) W_N^{pk} \sum_{r=0}^{N-1} f(r) W_N^{-rl} \right\} = \\ &= \sum_{p=0}^{N-1} \sum_{r=0}^{N-1} R_f(r-p) W_N^{(pk-rl)} = \sum_{p=0}^{N-1} \sum_{r=0}^{N-1} \sigma_f^2 \rho^{|r-p|} W_N^{(pk-rl)} = \\ &= \sigma_f^2 N \delta(k-l) \left[-\frac{\rho^{-1} W_N^k}{1-\rho^{-1} W_N^k} - \frac{1}{1-\rho^{-1} W_N^{-l}} \right] + \\ &+ \sigma_f^2 \left[\frac{1}{(1-\rho^{-1} W_N^k)(1-\rho W_N^{-l})} + \frac{1}{(1-\rho^{-1} W_N^{-l})(1-\rho W_N^k)} \right] \times (1-\rho), \end{aligned} \quad (2.44)$$

где $w_N = e^{-i \frac{2\pi}{N}}$ – фазовый множитель.

Первое слагаемое в (2.44) отлично от нуля только при $k = l$, то есть тогда, когда АКФ превращается в дисперсию. По сравнению с этой дисперсией при $N \gg 1$ вторым слагаемым можно пренебречь, то есть

$$R_F(k, l) \approx \sigma_f^2 N \delta(k - l) \left[-\frac{\rho^{-1} W_N^k}{1 - \rho^{-1} W_N^k} - \frac{1}{1 - \rho^{-1} W_N^{-l}} \right] \quad (2.45)$$

– единичный импульс с коэффициентом. То есть ДПФ сигнала является дискретным «почти» белым шумом. Положив в (2.45) $k = l = m$ определим его дисперсию в каждой точке:

$$\begin{aligned} \sigma_F^2(m) &= R_F(m, m) = \sigma_f^2 N \left[-\frac{\rho^{-1} W_N^m}{1 - \rho^{-1} W_N^m} - \frac{1}{1 - \rho^{-1} W_N^{-m}} \right] = \\ &= \sigma_f^2 N \frac{\rho^{-2} - 1}{1 + \rho^{-2} - \rho^{-1} (W_N^m + W_N^{-m})} = \sigma_f^2 N \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos \frac{2\pi}{N} m}. \end{aligned} \quad (2.46)$$

Видно, что дисперсия каждой спектральной компоненты $F(m)$ зависит от ее номера m , длины последовательности N и корреляционных свойств сигнала ρ .

Аналогичным путем можно вычислить и дисперсии ДПФ шума. Однако в нашем случае нет необходимости повторять весь ход преобразований. С учетом (2.42) можно, положив в (2.44) $\rho = 0$ и заменив индексы сигнала на индексы шума, сразу получить:

$$R_V(k, l) = N \sigma_v^2 \delta(k - l),$$

т.е. белый шум во временной области переходит в белый же шум в спектральной области. В отличие от дисперсии (2.46), дисперсия спектральных компонент шума не зависит от m :

$$\sigma_V^2(m) = \sigma_V^2 = N \sigma_v^2. \quad (2.47)$$

Таким образом, для модели наблюдения в спектральной области задача сводится к оценке белого шума с дисперсией (2.46) на фоне белого шума с дисперсией (2.47). Восстановление заключается в точечной оценке, то есть в умножении каждого спектрального отсчета на коэффициент λ_m :

$$\hat{F}(m) = \lambda_m G(m), \quad 0 \leq m \leq N-1, \quad (2.48)$$

где

$$\lambda_m = \frac{\sigma_F^2(m)}{\sigma_F^2(m) + \sigma_V^2} = \frac{d^2(1-\rho^2)}{d^2(1-\rho^2) + (1+\rho^2) - 2\rho \cos \frac{2\pi}{N}m}. \quad (2.49)$$

Далее полученная по (2.48) оценка $\hat{F}(m)$ переводится во временную область при помощи обратного ДПФ. Схема всей процедуры восстановления показана на рисунке 2.5. Такая процедура восстановления является асимптотически оптимальной при $N \rightarrow \infty$.

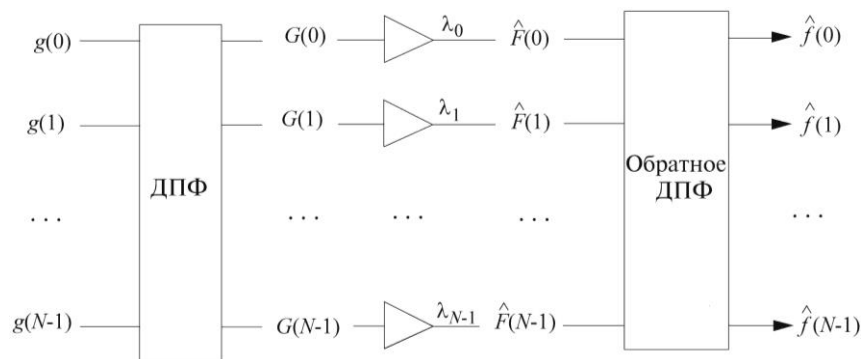


Рис. 2.5. Схема процедуры восстановления сигнала с использованием ДПФ

2.5 Восстановление сигнала физически-реализуемым БИХ фильтром

Рассмотренные выше способы реализации оптимального фильтра позволяют на практике воспользоваться результатами теории оптимального восстановления сигналов. Однако в некоторых случаях эти способы все же оказываются неприменимы. Например, при обработке сигналов в темпе процессов измерений не всегда имеется возможность обрабатывать его в «обратном» времени. Использование ДПФ при больших N оказывается для некоторых применений чрезмерно сложным, несмотря на наличие быстрых алгоритмов преобразования. В этой связи наряду с оптимальными фильтрами получили распространение *субоптимальные (квазиоптимальные) процедуры восстановления*, имеющие более простую реализацию.

Ниже рассмотрим восстановление сигнала БИХ-фильтром с импульсной характеристикой, удовлетворяющей условию:

$$h_{\text{восст}}(n) = 0 \text{ при } n < N_1. \quad (2.50)$$

Если отсчеты этой импульсной характеристики выбраны исходя из минимизации ошибки восстановления (2.7), то такой фильтр называется *физически реализуемым оптимальным фильтром* или, в случае независимости сигнала и шума в модели наблюдения, физически реализуемым фильтром Винера. Очевидно, что он будет работать несколько хуже, чем оптимальная восстанавливающая ЛИС-система, поскольку на часть отсчетов $h_{\text{восст}}(n)$ наложено сформулированное выше ограничение.

В зависимости от значения постоянной N_1 в условии (2.50) здесь возможны варианты, показанные на рисунке 2.6.

В первом случае (рисунок 2.6,а) осуществляется восстановление без задержки, оценка строится на основании наблюдаемого

сигнала в текущий и предыдущие моменты времени. Во втором случае (рисунок 2.6,б) восстановление производится только по «прошлым» значениям наблюдаемого (искаженного) сигнала, то есть имеет место предсказание (экстраполяция) сигнала. В третьем случае (рисунок 2.6,в) фильтр является уже физически нереализуемым, здесь восстановление производится с учетом некоторого числа “будущих” отсчетов сигнала. Такой фильтр может быть сведен к физически реализуемому, если допустить задержку в получении выходного сигнала. Здесь мы имеем процедуру восстановления с задержкой (интерполяции).

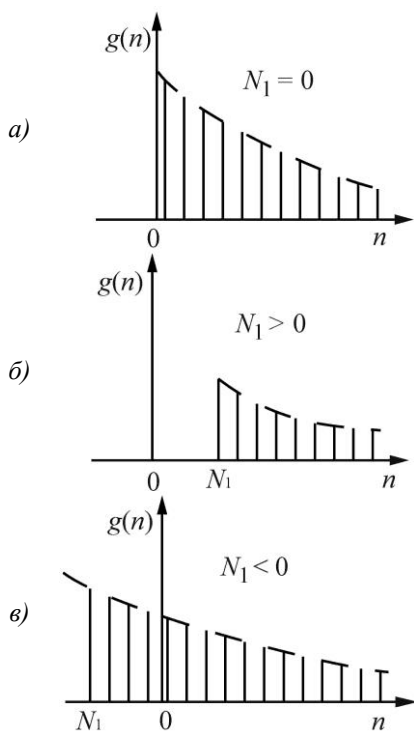


Рис. 2.6. Различные варианты ИХ квазиоптимального восстанавливающего БИХ-фильтра

Ниже мы ограничимся рассмотрением только первого случая.

Итак, построим физически реализуемый БИХ-фильтр, осуществляющий оптимальное (квазиоптимальное) восстановление сигнала без задержки. Для него с учетом (2.50) система уравнений (2.12) может быть переписана в конкретизированном виде:

$$\begin{cases} \sum_{k=0}^{\infty} h_{\text{восст}}(k) R_g(m-k) = R_{fg}(-m) & m \geq 0, \\ h_{\text{восст}}(m) = 0 & m < 0. \end{cases} \quad (2.51)$$

Рассмотрим сначала простой (но далекий от практики) случай, когда искаженный сигнал $g(n)$ является белым шумом с единичной дисперсией:

$$R_g(m) = \delta(m), \quad (2.52)$$

то есть представляет собой последовательность из некоррелированных отсчетов. Тогда, подставив (2.52) в (2.51), сразу получаем:

$$h_{\text{восст}}(m) = \begin{cases} R_{fg}(-m) & m \geq 0, \\ 0 & m < 0. \end{cases} \quad (2.53)$$

Соответственно, передаточная функция искомого фильтра:

$$H_{\text{восст}}(z) = \sum_{m=0}^{\infty} R_{fg}(-m) z^{-m}. \quad (2.54)$$

Здесь необходимо сделать небольшое отступление и ввести понятие (и обозначение) «реализуемой части последовательности» или ее z -преобразования. Для произвольной последовательности $\{f(n)\}_{n=-\infty}^{\infty}$ это:

$$[f(n)]_R = f(n)u(n), \quad (2.55)$$

$$[F(z)]_R = \sum_{n=-\infty}^{\infty} [f(n)]_R z^{-n} = \sum_{n=0}^{\infty} f(n) z^{-n}. \quad (2.56)$$

Если известно z -преобразование $F(z)$, то $[F(z)]_R$ можно найти двумя способами:

Первый способ универсален, но не всегда удобен в практическом применении. В соответствии с ним нужно взять обратное z -преобразование, полученную последовательность умножить на функцию единичного скачка (выполнить операцию (2.55)), и затем снова взять z -преобразование (выполнить (2.56)).

Второй способ специфичен для дробно-рациональных z -преобразований. Если $F(z)$ – такое z -преобразование, сходящееся в кольце

$$R_- < |z| < R_+,$$

то можно вычислить реализуемую часть, разложив его на простые дроби. Действительно, в общем случае (см. формулу (4.37) в [1]):

$$F(z) = \sum_k d_k z^{-k} + \sum_l \sum_{j=1}^{N_l} \frac{C_{jl}}{(1 - p_l z^{-1})^j},$$

где k, l – индексы, принадлежащие некоторым конечным множествам, по которым производится суммирование, p_l – полюсы кратности N_l . Реализуемая часть получится, если взять из этого разложения только те слагаемые, которые соответствуют последовательностям, отличным от нуля при $n \geq 0$:

$$[X(z)]_R = \sum_{k \geq 0} d_k z^{-k} + \sum_{l: |p_l| \leq R_-} \sum_{j=1}^{N_l} \frac{C_{jl}}{(1 - p_l z^{-1})^j},$$

– то есть часть дельта-импульсов и правосторонние экспоненты (для кратных полюсов – с полиномиальными множителями).

Итак, возвращаясь к задаче восстановления, если наблюдаемый сигнал $g(n)$ – белый шум с единичной дисперсией, то характеристики оптимального БИХ-фильтра, заданные соотношениями (2.53) и (2.54), могут быть записаны в форме:

$$h_{восст}(n) = \left[R_{fg}(-n) \right]_R, \quad (2.57)$$

$$H_{восст}(z) = \left[\Phi_{fg}(z^{-1}) \right]_R. \quad (2.58)$$

На практике, как уже говорилось, наблюдаемый сигнал $g(n)$ не является белым шумом. Но мы можем превратить его в таковой, пропустив через специально рассчитанную ЛИС-систему – “отбеливающий” фильтр. При этом вся процедура восстановления приобретает вид, показанный на рисунке 2.7, то есть будет представлять собой последовательное соединение отбеливающего и восстанавливающего фильтров.

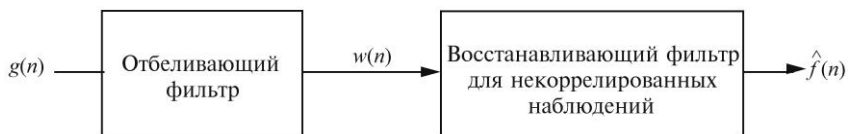


Рис. 2.7. Физически-реализуемый оптимальный восстанавливающий БИХ-фильтр

Пусть $H_{отб}(z)$ – передаточная функция отбеливающего фильтра, а $w(n)$ – последовательность на его выходе. Тогда:

$$H_{восст}(z) = H_{отб}(z) \left[\Phi_{fw}(z^{-1}) \right]_R. \quad (2.59)$$

Выразим входящие в (2.59) z -преобразования через известные нам характеристики. Во-первых, определим передаточную функцию отбеливающего фильтра. По смыслу его применения, должно быть:

$$\Phi_w(z) = H_{om\bar{o}}(z)H_{om\bar{o}}(z^{-1})\Phi_g(z) = 1$$

или

$$H_{om\bar{o}}(z)H_{om\bar{o}}(z^{-1}) = \frac{1}{\Phi_g(z)}. \quad (2.60)$$

Здесь необходимо решить задачу факторизации функции $\Phi_g(z)$ (см. п. 1.3), то есть представить ее в виде

$$\Phi_g(z) = \Phi_g^+(z)\Phi_g^+(z^{-1}), \quad (2.61)$$

где $\Phi_g^+(z)$ – передаточная функция физически реализуемой ЛИС–системы, имеющая нули, расположенные внутри единичной окружности. Последнее требование обусловлено тем, что из (2.60) и (2.61) мы получаем:

$$H_{om\bar{o}}(z) = \frac{1}{\Phi_g^+(z)}, \quad (2.62)$$

и упомянутые нули (ставшие теперь полюсами) обеспечат устойчивость отбеливающего фильтра.

Во-вторых, вычислим взаимный энергетический спектр последовательностей на входе и на выходе второго блока процедуры, показанной на рисунке 2.7. Можно показать, что

$$\Phi_{fw}(z) = \Phi_{fg}(z)H_{om\bar{o}}(z) = \frac{\Phi_{fg}(z)}{\Phi_g^+(z)}. \quad (2.63)$$

Подставив (2.62) и (2.63) в (2.59), окончательно получаем:

$$H_{\text{восст}}(z) = \frac{1}{\Phi_g^+(z)} \left[\frac{\Phi_{fg}(z^{-1})}{\Phi_g^+(z^{-1})} \right]_R. \quad (2.64)$$

Пример 2.3. Построим физически реализуемый восстанавливающий БИХ-фильтр для экспоненциально коррелированного сигнала, искаженного статистически независимым белым шумом, то есть для модели наблюдения (2.2б).

В силу независимости последовательностей в модели наблюдения имеем:

$$R_g(m) = R_f(m) + R_v(m).$$

Отсюда, переходя к z -преобразованиям, получаем:

$$\begin{aligned} \Phi_g(z) &= \Phi_f(z) + \Phi_v(z) = \\ &= \frac{\sigma_f^2(1-\rho^2)}{(1-\rho z^{-1})(1-\rho z)} + \sigma_v^2 = \sigma_v^2 \frac{\frac{1}{\alpha} + \alpha - (z + z^{-1})}{\frac{1}{\rho} + \rho - (z + z^{-1})}, \\ |\rho| < |z| < \frac{1}{|\rho|}, \end{aligned} \quad (2.65)$$

где значение параметра α определяется из уравнения

$$\frac{1}{\alpha} + \alpha = d^2 \left(\frac{1}{\rho} - \rho \right) + \left(\frac{1}{\rho} + \rho \right)$$

с условием $|\alpha| < 1$. Решение этого уравнения дает:

$$\alpha = \frac{1}{2\rho} \left\{ 1 + \rho^2 + d^2(1-\rho^2) \left[1 - \sqrt{\left(1 + \frac{1}{d^2}\right)^2 + \frac{4\rho^2}{d^2(1-\rho^2)}} \right] \right\}. \quad (2.66)$$

Формулу (2.65) можно представить в виде «симметричных» сомножителей:

$$\Phi_g(z) = \sigma_v^2 \frac{\rho}{\alpha} \frac{(1 - \alpha z^{-1})(1 - \alpha z)}{(1 - \rho z^{-1})(1 - \rho z)}, \quad |\rho| < |z| < \frac{1}{|\rho|},$$

после чего решение задачи факторизации становится очевидным:

$$\Phi_g^+(z) = \sqrt{\frac{\rho}{\alpha}} \sigma_v \frac{1 - \alpha z^{-1}}{1 - \rho z^{-1}}, \quad |\alpha| < 1, \quad |\rho| < 1.$$

Передаточная функция (2.62) отбеливающего фильтра здесь равна:

$$H_{отб}(z) = \frac{1}{\Phi_g^+(z)} = \frac{\sqrt{\alpha}}{\sqrt{\rho} \sigma_v} \cdot \frac{1 - \rho z^{-1}}{1 - \alpha z^{-1}}, \quad |z| > |\alpha|. \quad (2.67)$$

Далее определим взаимный энергетический спектр последовательностей:

$$\Phi_{fg}(z) = \Phi_f(z) = \frac{(1 - \rho^2) \sigma_f^2}{(1 - \rho z^{-1})(1 - \rho z)}, \quad |\rho| < |z| < \frac{1}{|\rho|},$$

и физически реализуемую часть z – преобразования, входящую как сомножитель в формулу (2.64):

$$\frac{\Phi_{fg}(z)}{\Phi_g^+(z)} = \frac{(1 - \rho^2) \sigma_f^2 \sqrt{\alpha}}{\sqrt{\rho} \sigma_v} \frac{1}{(1 - \rho z)(1 - \alpha z^{-1})}, \quad |\alpha| < |z| < \frac{1}{|\rho|};$$

$$\frac{\Phi_{fg}(z^{-1})}{\Phi_g^+(z^{-1})} = \frac{(1 - \rho^2) \sigma_f^2 \sqrt{\alpha}}{\sqrt{\rho} \sigma_v} \frac{1}{(1 - \rho z^{-1})(1 - \alpha z)}, \quad |\rho| < |z| < \frac{1}{|\alpha|}. \quad (2.68)$$

Для отыскания реализуемой части разложим z -преобразование (2.68) на простые дроби (при этом коэффициент пока оставим в

стороне), а область сходимости будем указывать непосредственно под формулами:

$$\frac{1}{(1-\rho z^{-1})(1-\alpha z)} = \frac{1}{1-\alpha\rho} \times \frac{1}{1-\rho z^{-1}} \times \frac{1}{1-\alpha\rho} \frac{1}{1-\frac{1}{\alpha}z^{-1}} .$$

$$\begin{array}{ccc} |\rho| < |z| < \frac{1}{|\alpha|} & |z| > |\rho| & |z| < \frac{1}{|\alpha|} \end{array}$$

Реализуемой части соответствует первое слагаемое, то есть с учетом коэффициента из (2.68) имеем:

$$\left[\frac{\Phi_{fg}(z^{-1})}{\Phi_g^+(z^{-1})} \right]_R = \frac{(1-\rho^2)\sigma_f^2\sqrt{\alpha}}{\sqrt{\rho}\sigma_v(1-\alpha\rho)} \frac{1}{1-\rho z^{-1}} . \quad (2.69)$$

Передаточная функция всего восстанавливающего фильтра получается подстановкой (2.67) и (2.69) в (2.64):

$$H_{восст}(z) = \frac{\sqrt{\alpha}}{\sqrt{\rho}\sigma_v} \frac{1-\rho z^{-1}}{1-\alpha z^{-1}} \frac{(1-\rho^2)\sigma_f^2\sqrt{\alpha}}{\sqrt{\rho}\sigma_v(1-\alpha\rho)} \frac{1}{1-\rho z^{-1}} =$$

$$= \frac{C}{1-\alpha z^{-1}}, \quad (2.70)$$

где на последнем шаге обозначено $C = \frac{(1-\rho^2)\alpha d^2}{\rho(1-\alpha\rho)}$, α – определяется соотношением (2.66). Из (2.70) легко строится разностное уравнение, задающее рекуррентный алгоритм фильтрации сигнала:

$$\tilde{f}(n) = \alpha \tilde{f}(n-1) + C g(n) .$$

Простая структурная схема ЛИС–системы, соответствующей полученному уравнению, показана на рисунке 2.8.

Импульсная характеристика восстанавливающего БИХ-фильтра:

$$h_{\text{восст}}(n) = C \alpha^n u(n).$$

В соответствии с формулой (2.16) ошибка восстановления сигнала здесь равна:

$$\bar{\varepsilon}^2 = \sigma_v^2 h_{\text{восст}}(0) = \sigma_v^2 C = \frac{\sigma_v^2 (1 - \rho^2) \alpha d^2}{\rho(1 - \alpha \rho)}.$$

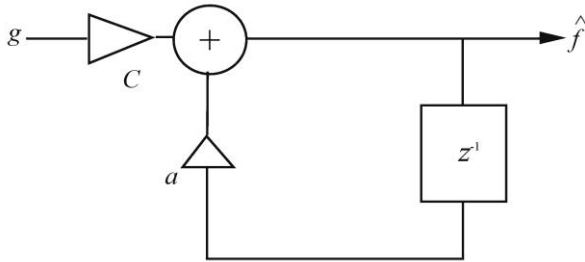


Рис. 2.8. Структурная схема ЛИС-системы, реализующей рекуррентное восстановление сигнала

2.6 Восстановление сигнала КИХ-фильтром

Построим теперь субоптимальный восстанавливающий КИХ-фильтр. В этом случае за оценку сигнала $\hat{f}(n)$ принимается взвешенная сумма конечного числа отсчетов наблюдаемого сигнала $g(n)$, то есть здесь оценка строится не рекурсивно, как результат непосредственного вычисления свертки

$$\hat{f}(n) = \sum_{k \in D} h_{\text{восст}}(k) g(n - k), \quad (2.71)$$

где D – конечное множество отсчетов, задающее «окно» обработки.

Выбрав область D вокруг восстанавливаемого отсчета достаточно большого размера и рассчитав оптимальные коэффициенты

КИХ-фильтра, можно получить среднеквадратичную погрешность восстановления, очень близкую к минимально достижимой, обеспечиваемой оптимальным физически нереализуемым линейным фильтром. Более того, даже при относительно небольших размерах окна обработки ошибка получается, как правило, меньше, чем у физически реализуемого восстанавливающего БИХ-фильтра. Это происходит благодаря тому, что в данном случае формируется «двусторонняя» (интерполяционная) оценка, в которой учтены не только «прошлые», но и некоторое число «будущих» отсчетов наблюдаемого сигнала. Естественно, в этом случае восстановление реализуется с некоторой задержкой.

Задача синтеза субоптимального восстанавливающего КИХ-фильтра заключается в определении значений $h_{восст}(n)$ в пределах окна обработки, обеспечивающих минимум среднеквадратичной ошибки восстановления. Как и раньше, они определяются из системы уравнений (2.12). Отличие от предыдущих случаев состоит в том, что теперь область D содержит конечное число элементов – N_D . Поэтому уравнение Винера-Хопфа (первая строка системы (2.12)) определяет систему из N_D линейных алгебраических уравнений с таким же числом неизвестных – значений искомой импульсной характеристики. Методы решений таких систем хорошо известны.

Пример 2.4. Построим простейший КИХ-фильтр вида (2.71) – процедуру восстановления сигнала по трем точкам:

$$\begin{aligned} \tilde{f}(n) &= \sum_{k=-1}^1 h_{восст}(k) g(n-k) = \\ &= h_{восст}(-1) g(n+1) + h_{восст}(0) g(n) + h_{восст}(1) g(n-1) \end{aligned}$$

для экспоненциально коррелированного сигнала, искаженного статически независимым белым шумом (для модели наблюде-

ния (2.26). Здесь $D = \{-1, 0, 1\}$. Из уравнения Винера-Хопфа получаем:

$$\begin{cases} R_g(0)h_{\text{восст}}(-1) + R_g(-1)h_{\text{восст}}(0) + R_g(-2)h_{\text{восст}}(1) = R_{fg}(1), \\ R_g(1)h_{\text{восст}}(-1) + R_g(0)h_{\text{восст}}(0) + R_g(-1)h_{\text{восст}}(1) = R_{fg}(0), \\ R_g(2)h_{\text{восст}}(-1) + R_g(1)h_{\text{восст}}(0) + R_g(0)h_{\text{восст}}(1) = R_{fg}(-1). \end{cases}$$

В данном случае

$$\begin{aligned} R_g(m) &= R_f(m) + R_v(m) = \sigma_f^2 \rho^{|m|} + \sigma_v^2 \delta(m), \\ R_{fg}(-m) &= R_f(-m) = R_f(m) = \sigma_f^2 \rho^{|m|}, \end{aligned}$$

поэтому записанная система уравнений конкретизируется:

$$\begin{cases} (\sigma_f^2 + \sigma_v^2)h_{\text{восст}}(-1) + \sigma_f^2 \rho h_{\text{восст}}(0) + \sigma_f^2 \rho^2 h_{\text{восст}}(1) = \sigma_f^2 \rho, \\ \sigma_f^2 \rho h_{\text{восст}}(-1) + (\sigma_f^2 + \sigma_v^2)h_{\text{восст}}(0) + \sigma_f^2 \rho h_{\text{восст}}(1) = \sigma_f^2, \\ \sigma_f^2 \rho^2 h_{\text{восст}}(1) + \sigma_f^2 \rho h_{\text{восст}}(0) + (\sigma_f^2 + \sigma_v^2)h_{\text{восст}}(1) = \sigma_f^2 \rho. \end{cases} \quad (2.72)$$

Решение системы (2.72) имеет вид:

$$\begin{aligned} h_{\text{восст}}(0) &= \frac{1 + \frac{1}{d^2} - \rho^2}{\left(1 + \frac{1}{d^2}\right)\left(1 + \frac{1}{d^2} + \rho^2\right) - 2\rho^2}, \\ h_{\text{восст}}(-1) = h_{\text{восст}}(1) &= \frac{\frac{1}{d^2} \rho}{\left(1 + \frac{1}{d^2}\right)\left(1 + \frac{1}{d^2} + \rho^2\right) - 2\rho^2}, \end{aligned}$$

где, как и раньше, использовано обозначение $d^2 = \frac{\sigma_f^2}{\sigma_v^2}$.

Полученный КИХ-фильтр может быть реализован с задержкой на один шаг в форме прямой свертки так, как показано на рисунке 2.9.

Ошибка восстановления сигнала здесь опять определяется по формуле (2.16):

$$\bar{\varepsilon}^2 = \sigma_v^2 h_{восст}(0) = \sigma_v^2 \frac{1 + \frac{1}{d^2} - \rho^2}{\left(1 + \frac{1}{d^2}\right) \left(1 + \frac{1}{d^2} + \rho^2\right) - 2\rho^2}.$$

Достоинство нерекурсивных процедур восстановления состоит в простоте их расчета. Для построения восстанавливающего КИХ-фильтра достаточно решить систему линейных уравнений, а не решать сложную задачу факторизации энергетических спектров. Кроме того, как уже отмечалось, КИХ-фильтр может обеспечить качество восстановления более близкое к оптимальному, чем физически реализуемый винеровский фильтр.

Еще одно достоинство заключается в том, что данная методика расчета процедур восстановления легко обобщается на случай обработки двумерных сигналов.

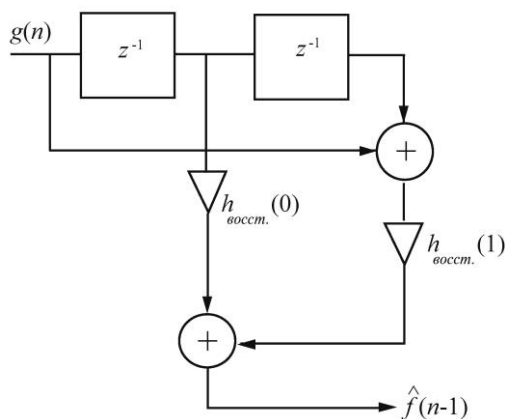


Рис. 2.9. КИХ-фильтр, реализованный в форме прямой свертки с задержками

3 ДВУМЕРНЫЕ СЛУЧАЙНЫЕ СИГНАЛЫ И СИСТЕМЫ

3.1 Двумерные случайные последовательности, их характеристики и прохождение через ЛИС-системы

Так же как и в одномерном случае под двумерной случайной последовательностью или случайным полем будем понимать последовательность, элементы которой являются случайными величинами. Остановимся кратко на характеристиках двумерных случайных последовательностей.

Вместо термина «стационарный процесс» в двумерном случае иногда используется термин «однородное случайное поле», АКФ которого зависит от двух аргументов:

$$R_f(t_1, t_2, \tau_1, \tau_2) = R_f(t_1 - \tau_1, t_2 - \tau_2).$$

Выражения для оценок математического ожидания и АКФ поля с непрерывными аргументами, наблюдаемого на двумерном интервале величиной T_1, T_2 , выглядят следующим образом:

$$\hat{\mu}_f \approx \frac{1}{T_1 T_2} \int_0^{T_1} \int_0^{T_2} f(t_1, t_2) dt_1 dt_2,$$
$$\hat{R}_f(\tau_1, \tau_2) \approx \frac{1}{T_1 T_2} \int_0^{T_1} \int_0^{T_2} (f(t_1, t_2) - \mu_f)(f(t_1 + \tau_1, t_2 + \tau_2) - \mu_f) dt_1 dt_2.$$

Оценки числовых характеристик случайной последовательности в двумерном случае так же определяются с использованием свойства эргодичности и имеют следующий вид для интервала наблюдения ($1 \leq n_1 \leq N_1, 1 \leq n_2 \leq N_2$):

$$\hat{\mu}_f \approx \frac{1}{N_1 N_2} \sum_{k_1=1}^{N_1} \sum_{k_2=1}^{N_2} f(k_1, k_2), \quad (3.1)$$

$$\hat{R}_f(n_1, n_2) \approx \frac{1}{(N_1 - n_1)(N_2 - n_2)} \times \\ \times \sum_{k_1=1}^{N_1 - n_1} \sum_{k_2=1}^{N_2 - n_2} (f(k_1, k_2) - \mu_f)(f(n_1 + k_1, n_2 + k_2) - \mu_f). \quad (3.2)$$

Связь АКФ и спектральной плотности мощности (энергетического спектра) двумерной стационарной последовательности определяется уравнением:

$$\Phi_f(e^{i\omega_1}, e^{i\omega_2}) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} R_f(n_1, n_2) e^{i\omega_1 n_1 - i\omega_2 n_2}, \quad (3.3)$$

$$R_f(n_1, n_2) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Phi_f(e^{i\omega_1} e^{i\omega_2}) e^{i\omega_1 n_1 + i\omega_2 n_2} d\omega_1 d\omega_2. \quad (3.4)$$

Прохождение двумерной последовательности через ЛИС-систему описывается теми же соотношениями, что и в одномерном случае, с обобщением на две координаты.

Пусть известны характеристики входного сигнала – двумерной стационарной случайной последовательности $f(n_1, n_2)$: среднее значение μ_f , АКФ R_f и энергетический спектр Φ_f . Требуется получить соответствующие характеристики для последовательности $g(n_1, n_2)$ на выходе устойчивой ЛИС-системы с импульсной характеристикой $h(n_1, n_2)$, а также взаимные статистические характеристики входной и выходной последовательностей.

Среднее значение для выходной последовательности с учетом стационарности сигнала и известной формулы свертки определяется следующим образом:

$$\begin{aligned}
\mu_g &= E\{g(n_1, n_2)\} = \\
&= E\left\{\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} h(k_1, k_2) f(n_1 - k_1, n_2 - k_2)\right\} = \\
&= \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} h(k_1, k_2) E\{f(n_1 - k_1, n_2 - k_2)\} = \\
&= \mu_f \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} h(k_1, k_2).
\end{aligned} \tag{3.5}$$

Если ЛИС-система описана не импульсной характеристикой, а частотной $H(e^{i\omega_1}, e^{i\omega_2})$ или передаточной функцией $H(z_1, z_2)$, то для вычисления среднего значения выходной последовательности можно воспользоваться соотношениями:

$$\mu_g = \mu_f H(e^{i\omega_1}, e^{i\omega_2}) \Big|_{\omega_1=\omega_2=0}, \quad \mu_g = \mu_f H(z_1, z_2) \Big|_{z_1=z_2=1}. \tag{3.6}$$

В дальнейшем для сокращения изложения будем полагать $\mu_g = \mu_f = 0$. При невыполнении этого равенства всегда можно учесть математическое ожидание и его преобразование отдельно на основании формул (3.5) и (3.6).

АКФ выходной последовательности

$$R_g(n_1, n_2) = E\left\{\left(g(k_1, k_2) - \mu_g\right)\left(g(k_1 + n_1, k_2 + n_2) - \mu_g\right)\right\}$$

определяется следующим образом:

$$R_g(n_1, n_2) = \sum_{k=-\infty}^{\infty} \left[\sum_{l=-\infty}^{\infty} h(l_1, l_2) h(l_1 + k_1, l_2 + k_2) \right] R_f(n_1 - k_1, n_2 - k_2).$$

Последнее выражение, записанное с использованием оператора свертки, выглядит следующим образом:

$$R_g(n_1, n_2) = h(n_1, n_2) * h(-n_1, -n_2) * R_f(n_1, n_2). \quad (3.7)$$

ВКФ входной и выходной последовательностей вычисляется в виде

$$\begin{aligned} R_{fg}(n_1, n_2) &= E \left\{ \left(f(n_1, n_2) - \mu_f \right) \left(g(n_1 + k_1, n_2 + k_2) - \mu_g \right) \right\} = \\ &= \sum_{k=-\infty}^{\infty} h(k_1, k_2) R_f(n_1 - k_1, n_2 - k_2) = \\ &= h(n_1, n_2) * R_f(n_1, n_2). \end{aligned} \quad (3.8)$$

То есть искомая характеристика является сверткой импульсной характеристики ЛИС-системы и АКФ входного сигнала.

Энергетический спектр двумерной стационарной последовательности на выходе ЛИС-системы легко выводится из уже полученного соотношения (3.7). Действительно, с учетом свойств z -преобразования (см. п. 6.8 учебного пособия [1]) имеем

$$\Phi_g(z_1, z_2) = H(z_1, z_2) H(z_1^{-1}, z_2^{-1}) \Phi_f(z_1, z_2), \quad (3.9)$$

и далее, положив $z_1 = e^{i\omega_1}$, $z_2 = e^{i\omega_2}$, получаем собственно энергетический спектр:

$$\Phi_g(e^{i\omega_1}, e^{i\omega_2}) = H(e^{i\omega_1}, e^{i\omega_2}) H(e^{-i\omega_1}, e^{-i\omega_2}) \Phi_f(e^{i\omega_1}, e^{i\omega_2}). \quad (3.10)$$

Частотная характеристика обладает известной симметрией и выражение (3.10) может быть записано в более компактной форме:

$$\Phi_g(e^{i\omega_1}, e^{i\omega_2}) = \left| H(e^{i\omega_1}, e^{i\omega_2}) \right|^2 \Phi_f(e^{i\omega_1}, e^{i\omega_2}). \quad (3.11)$$

Получаем, что энергетический спектр последовательности на выходе ЛИС-системы равен энергетическому спектру входной последовательности, умноженному на квадрат модуля частотной характеристики системы.

Взаимный энергетический спектр входной и выходной последовательности вычисляется аналогично из (3.8):

$$\Phi_{fg}(z_1, z_2) = H(z_1, z_2)\Phi_f(z_1, z_2), \quad (3.12)$$

и далее при $z_1 = e^{i\omega_1}, z_2 = e^{i\omega_2}$:

$$\Phi_{fg}(e^{i\omega_1}, e^{i\omega_2}) = H(e^{i\omega_1}, e^{i\omega_2})\Phi_f(e^{i\omega_1}, e^{i\omega_2}) \quad (3.13)$$

он равен произведению частотной характеристики системы и энергетического спектра входной последовательности.

Общего подхода к факторизации энергетического спектра в двумерном случае не существует.

3.2 Двумерная оптимальная линейная фильтрация

На двумерный случай переносятся все основные результаты теории оптимальной линейной фильтрации одномерных сигналов.

Пусть имеется линейная дискретная модель наблюдения двумерного сигнала:

$$g(n_1, n_2) = f(n_1, n_2) ** h(n_1, n_2) + v(n_1, n_2), \quad (3.14)$$

где $h(n_1, n_2)$ – импульсная характеристика искажающей двумерной ЛИС-системы; $f(n_1, n_2)$ – полезный сигнал, стационарное случайное поле; $v(n_1, n_2)$ – помеха, тоже стационарное случайное поле.

Пусть восстановление двумерного сигнала осуществляется при помощи ЛИС-системы с импульсной характеристикой $h_{восст}(n_1, n_2)$:

$$\hat{f}(n_1, n_2) = g(n_1, n_2) ** h_{восст}(n_1, n_2), \quad (3.15)$$

которая отлична от нуля только в некоторой двумерной области D :

$$h_{\text{восст}}(n_1, n_2) = 0 \text{ для всех } (n_1, n_2) \notin D.$$

Требуется найти восстанавливающую систему, которая при сформулированных ограничениях на импульсную характеристику обеспечивает минимум среднеквадратичной ошибки восстановления:

$$E \left\{ \left[\hat{f}(n_1, n_2) - f(n_1, n_2) \right]^2 \right\} \rightarrow \min.$$

Параметры восстанавливающей ЛИС-системы (значения ее импульсной характеристики) определяются из двумерного аналога системы уравнений (2.12):

$$\begin{cases} \sum_{(k_1, k_2) \in D} h_{\text{восст}}(k_1, k_2) R_g(n_1 - k_1, n_2 - k_2) = \\ \qquad \qquad \qquad = R_{fg}(-n_1, -n_2), \quad (n_1, n_2) \in D; \\ h_{\text{восст}}(n_1, n_2) = 0, \quad (n_1, n_2) \notin D. \end{cases} \quad (3.16)$$

Погрешность восстановления, осуществленного ЛИС-системой с параметрами, определенными из (3.16), задается выражением:

$$\varepsilon^2 = \sigma_f^2 - \sum_{(k_1, k_2) \in D} \sum h_{\text{восст}}(k_1, k_2) R_{fg}(-k_1, -k_2). \quad (3.17)$$

Передаточная функция оптимальной восстанавливающей ЛИС-системы, если область D не ограничена по обоим аргументам, имеет вид (см. также формулу (2.18)):

$$H_{\text{восст}}(z_1, z_2) = \frac{\Phi_{fg}(z_1^{-1}, z_2^{-1})}{\Phi_g(z_1, z_2)}. \quad (3.18)$$

При этом в случае, когда сигнал и шум статистически независимы, передаточная функция оптимальной восстанавливающей

системы (двумерного фильтра Винера-Колмогорова) определяется аналогично формуле (2.24):

$$H_{восст}(z_1, z_2) = \frac{H(z_1^{-1}, z_2^{-1}) \cdot \Phi_f(z_1, z_2)}{H(z_1^{-1}, z_2^{-1}) \cdot H(z_1, z_2) \cdot \Phi_f(z_1, z_2) + \Phi_v(z_1, z_2)}. \quad (3.19)$$

Энергетический спектр (точнее, z-преобразование АКФ) остаточной ошибки восстановления сигнала фильтром Винера задается выражением (см. также (2.25)):

$$\Phi_\varepsilon(z_1, z_2) = \frac{\Phi_v(z_1, z_2) \Phi_f(z_1, z_2)}{H(z_1, z_2) H(z_1^{-1}, z_2^{-1}) \Phi_f(z_1, z_2) + \Phi_v(z_1, z_2)}. \quad (3.20)$$

Принципиальным отличием двумерного случая от одномерного является невозможность выполнения факторизации передаточной функции. Поэтому обычно решение задачи двумерной фильтрации осуществляется в спектральной области. Такая процедура включает в себя три этапа.

На *первом шаге* обработки матрица отсчетов наблюдаемого поля $g(n_1, n_2)$ ($0 \leq n_1 \leq N_1 - 1; 0 \leq n_2 \leq N_2 - 1$) подвергается двумерному ДПФ, в результате чего вычисляется дискретный спектр (трансформанты) – $G(k_1, k_2)$ ($0 \leq k_1 \leq N_1 - 1; 0 \leq k_2 \leq N_2 - 1$). Эти трансформанты при достаточно больших N_1, N_2 оказываются практически некоррелированными, и поэтому оптимальной является их точечная оценка.

Процедура точечного оценивания – *второй шаг* обработки:

$$\hat{F}(k_1, k_2) = \lambda_{k_1, k_2} \times G(k_1, k_2), \quad (3.21)$$

где λ_{k_1, k_2} – коэффициенты точечной оценки; $\hat{F}(k_1, k_2)$ – трансформанты оценки полезного сигнала.

На *третьем шаге* при помощи обратного двумерного ДПФ переходим от трансформант $\hat{F}(k_1, k_2)$ в пространственную область, то есть получаем искомую оценку полезного сигнала – $\hat{f}(n_1, n_2)$.

Для получения коэффициентов точечной оценки воспользуемся следующим подходом. Процедура оптимального восстановления, описываемая сверткой (3.15), в z-преобразованиях записывается в виде:

$$\hat{F}(z_1, z_2) = G(z_1, z_2) \cdot H_{восст}(z_1, z_2). \quad (3.22)$$

Положив $z_1 = e^{i\omega_1}$, $z_2 = e^{i\omega_2}$ можно записать аналогичное соотношение для спектров и частотной характеристики оптимального фильтра

$$\hat{F}(e^{i\omega_1}, e^{i\omega_2}) = G(e^{i\omega_1}, e^{i\omega_2}) \cdot H_{восст}(e^{i\omega_1}, e^{i\omega_2}).$$

Восстановленный сигнал $\hat{f}(n_1, n_2)$ в используемом подходе к реализации восстанавливающего фильтра через ДПФ рассматривается на ограниченном двумерном интервале ($0 \leq n_1 \leq N_1 - 1$; $0 \leq n_2 \leq N_2 - 1$), т.е. это двумерная *последовательность конечной длины*. Для такой последовательности известна связь непрерывного спектра с ДПФ (см. параграфы 4.3 и 6.6 учебного пособия [1]):

$$\hat{F}(k_1, k_2) = \hat{F}(e^{i\omega_1}, e^{i\omega_2}) \left| \begin{array}{l} \omega_1 = \frac{2\pi k_1}{N_1}, 0 \leq k_1 \leq N_1 - 1 \\ \omega_2 = \frac{2\pi k_2}{N_2}, 0 \leq k_2 \leq N_2 - 1 \end{array} \right.$$

То есть, взяв дискретный ряд частот, можно записать:

$$\hat{F}(k_1, k_2) = G(k_1, k_2) \cdot H_{восст}(k_1, k_2), \quad (3.23)$$

где $G(k_1, k_2)$ – ДПФ сигнала; $H_{восст}(k_1, k_2)$ – отсчеты частотной характеристики оптимального фильтра.

Строго говоря, эти отсчеты не совсем являются ДПФ импульсной характеристики $h_{восст}(n_1, n_2)$, получаемой из решения (3.16), которая в общем случае оказывается последовательностью бесконечной длины. Однако, при $N_1, N_2 \gg 1$ это несоответствие оказывается небольшим, а решение – близким к оптимальному.

Сравнивая выражение (3.23) с (3.21), видим, что коэффициенты точечной оценки

$$\begin{aligned} \lambda_{k_1, k_2} &= H(k_1, k_2) = \\ &= H_{восст} \left(e^{i\omega_1}, e^{i\omega_2} \right) \left| \begin{array}{l} \omega_1 = \frac{2\pi k_1}{N_1}, 0 \leq k_1 \leq N_1 - 1 \\ \omega_2 = \frac{2\pi k_2}{N_2}, 0 \leq k_2 \leq N_2 - 1 \end{array} \right. \end{aligned} \quad (3.24)$$

или

$$\begin{aligned} \lambda_{k_1, k_2} &= H(k_1, k_2) = \\ &= H_{восст} (z_1, z_2) \left| \begin{array}{l} z_1 = e^{i\frac{2\pi k_1}{N_1}}, 0 \leq k_1 \leq N_1 - 1 \\ z_2 = e^{i\frac{2\pi k_2}{N_2}}, 0 \leq k_2 \leq N_2 - 1. \end{array} \right. \end{aligned} \quad (3.25)$$

Пример 3.1. Пусть имеется упрощенная модель наблюдения

$$g(n_1, n_2) = f(n_1, n_2) + v(n_1, n_2), \quad (3.26)$$

где $f(n_1, n_2)$ – стационарное случайное поле с АКФ

$$R_f(p_1, p_2) = \sigma_f^2 \rho^{|p_1| + |p_2|}, \quad (3.27)$$

$v(n_1, n_2)$ – стационарный дискретный белый шум, статистически независимый от сигнала.

$$R_v(p_1, p_2) = \sigma_v^2 \delta(p_1, p_2). \quad (3.28)$$

Определим коэффициенты λ_{k_1, k_2} для двумерного оптимального фильтра.

Энергетические спектры сигнала и шума:

$$\Phi_f(z_1, z_2) = \frac{\sigma_f^2 (1 - \rho^2)^2}{\left[1 + \rho^2 - \rho(z_1 + z_1^{-1})\right] \left[1 + \rho^2 - \rho(z_2 + z_2^{-1})\right]}; \quad (3.29)$$

$$\Phi_v(z_1, z_2) = \sigma_v^2. \quad (3.30)$$

Передаточная функция оптимального фильтра получается из (3.20) при $H(z_1, z_2) = 1$:

$$\begin{aligned} H_{восст}(z_1, z_2) &= \frac{\Phi_f(z_1, z_2)}{\Phi_f(z_1, z_2) + \Phi_v(z_1, z_2)} = \\ &= \frac{\sigma_f^2 (1 - \rho^2)^2}{\sigma_f^2 (1 - \rho^2)^2 + \sigma_v^2 \left[1 + \rho^2 - \rho(z_1 + z_1^{-1})\right] \left[1 + \rho^2 - \rho(z_2 + z_2^{-1})\right]} = \\ &= \frac{d^2 (1 - \rho^2)^2}{d^2 (1 - \rho^2)^2 + \left[1 + \rho^2 - \rho(z_1 + z_1^{-1})\right] \left[1 + \rho^2 - \rho(z_2 + z_2^{-1})\right]}, \quad (3.31) \end{aligned}$$

где $d^2 = \frac{\sigma_f^2}{\sigma_v^2}$.

Коэффициенты оптимального восстанавливающего фильтра определяются соотношением:

$$\lambda_{k_1, k_2} = H_{восст}(z_1, z_2) \begin{cases} z_1 = e^{i \frac{2\pi k_1}{N_1}}, 0 \leq k_1 \leq N_1 - 1 \\ z_2 = e^{i \frac{2\pi k_2}{N_2}}, 0 \leq k_2 \leq N_2 - 1 \end{cases} =$$

$$= \frac{d^2(1-\rho^2)^2}{d^2(1-\rho^2)^2 + \left[1 + \rho^2 - 2\rho \cos \frac{2\pi k_1}{N_1}\right] \left[1 + \rho^2 - 2\rho \cos \frac{2\pi k_2}{N_2}\right]}. \quad (3.32)$$

Основным достоинством алгоритма восстановления сигнала с помощью ДПФ является его *универсальность*, т.е. применимость для любых линейных моделей наблюдения. Этот алгоритм является асимптотически оптимальным с ростом размера обрабатываемых матриц - прямоугольных фрагментов двумерного сигнала.

Очень серьезный недостаток алгоритма – высокие требования к объему оперативной памяти компьютера, трудоемкость и невозможность обработки в темпе поступления информации.

Однако двумерность обрабатываемых сигналов дает возможность преодолеть некоторые из указанных недостатков при построении *спектрально-рекуррентного алгоритма восстановления*. Общая схема спектрально-рекуррентного алгоритма такова.

Пусть в матрице $g(n_1, n_2)$ первый индекс $n_1 \in [0, N_1 - 1]$ означает номер строки, а второй – $n_2 \in [0, N_2 - 1]$ – номер отсчета в строке.

Сначала вычисляется одномерное ДПФ для строк матрицы наблюдаемого (т.е. искаженного) поля:

$$g^*(n_1, k_2) = \sum_{n_2=0}^{N_2-1} g(n_1, n_2) \cdot e^{-i \frac{2\pi}{N_2} k_2 n_2}, \quad 0 \leq k_2 \leq N_2 - 1. \quad (3.33)$$

При этом получается *дискретный полуспектр* $g^*(n_1, k_2)$ – двумерная последовательность, один из аргументов которой (n_1) соответствует пространственной координате, а другой (k_2) – частотной. В силу декоррелирующего свойства ДПФ при $N_2 \gg 1$ – элементы полуспектра в каждой строке будут практически независимы друг от друга. Это означает, что поле в полуспектральной области распадается на N_2 независимых последовательностей, соответствующих столбцам. Полуспектр $g^*(n_1, k_2)$ можно рассматривать как совокупность одномерных сигналов с аргументом n_1 , а k_2 – принимает смысл просто параметра, порядкового номера последовательности, т.е. можно его переобозначить: $g_k^*(n)$. Для того, чтобы получить оценку полезного сигнала, осуществляется *одномерная* оптимальная фильтрация каждого столбца полуспектра (для этого строится двухпроходный алгоритм). После такой фильтрации получаем оценку в полуспектральной области – $\hat{f}^*(n_1, k_2)$. На заключительном этапе эта оценка переводится в пространственную область при помощи одномерного обратного ДПФ, выполняемого по строкам:

$$\hat{f}(n_1, n_2) = \frac{1}{N_2} \sum_{k_2=0}^{N_2-1} \hat{f}^*(n_1, k_2) \cdot e^{i \frac{2\pi}{N_2} k_2 n_2}, \quad 0 \leq n_2 \leq N_2 - 1. \quad (3.34)$$

Рассмотрим, как рассчитывается алгоритм восстановления в полуспектральной области. Опять воспользуемся соотношением

$$(3.22). \text{ Положим в нем } z_2 = e^{i \frac{2\pi}{N_2} k_2}, \quad 0 \leq k_2 \leq N_2 - 1. \text{ Тогда}$$

$$\hat{F} \begin{pmatrix} z_1, e^{i \frac{2\pi}{N_2} k_2} \end{pmatrix} = G \begin{pmatrix} z_1, e^{i \frac{2\pi}{N_2} k_2} \end{pmatrix} H_{\text{восст}} \begin{pmatrix} z_1, e^{i \frac{2\pi}{N_2} k_2} \end{pmatrix}.$$

Или, используя обозначения для дискретного полуспектра:

$$\hat{F}^*(z_1, k_2) = G^*(z_1, k_2) H_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right). \quad (3.35)$$

Здесь k_2 является параметром. Задача состоит в построении алгоритма фильтрации одномерного сигнала, т.е. в построении ЛИС-

системы с передаточной функцией $H_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right)$. Здесь уже

можно произвести факторизацию:

$$H_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right) = H^+_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right) \cdot H^-_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right), \quad (3.36)$$

где $H^+_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right)$ – передаточная функция ЛИС-системы,

реализуемой в «прямом времени», т.е. при движении в направлении увеличения аргумента n_1 (вниз по столбцам);

$H^-_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right)$ – передаточная функция ЛИС-системы, реа-

лизуемой в «обратном времени», т.е. при движении вверх по столбцам.

Пример 3.2. Пусть имеется модель наблюдения (3.26), где полезный сигнал – стационарное поле с биэкспоненциальной изотропной АКФ (3.27), шум – стационарный белый, статистически независимый от сигнала, с дисперсией σ_v^2 .

Построим двухпроходную процедуру оптимального оценивания в полуспектральной области. Передаточная функция оптимального фильтра (см. также (3.31)):

$$H_{\text{восст}}(z_1, z_2) = \frac{\Phi_f(z_1, z_2)}{\Phi_f(z_1, z_2) + \Phi_v(z_1, z_2)} = \frac{d^2(1-\rho^2)^2}{d^2(1-\rho^2)^2 + \left[1 + \rho^2 - \rho(z_1 + z_1^{-1})\right] \left[1 + \rho^2 - \rho(z_2 + z_2^{-1})\right]}. \quad (3.37)$$

Положим $z_2 = e^{i\frac{2\pi}{N_2}k_2}$, обозначим $z_1 + z_1^{-1} = t$ и сократим дробь на $d^2(1-\rho^2)^2$. Тогда

$$H_{\text{восст}} \left(t, e^{i\frac{2\pi}{N_2}k_2} \right) = \frac{1}{1 + d^2 - \frac{2\rho d^2}{(1-\rho^2)} \cos \frac{2\pi k_2}{N_2} - \frac{\rho d^2}{(1-\rho^2)} t + \frac{2\rho^2 d^2}{(1-\rho^2)^2} t \cos \frac{2\pi k_2}{N_2}} = \frac{1}{B(k_2)/A - t C(k_2)/A}. \quad (3.38)$$

где $A = d^2(1-\rho^2)^2$,

$$B(k_2) = d^2(1-\rho^2) + (1-\rho^2)^2 - 2\rho(1+\rho^2) \cos \frac{2\pi}{N_2} k_2,$$

$$C(k_2) = \rho(1-\rho^2) - 2\rho^2 \cos \frac{2\pi}{N_2} k_2.$$

Полученное выражение для одномерной передаточной функции всегда удается факторизовать, т.е. представить в виде произведения:

$$\begin{aligned}
 H_{\text{восст}} \left(z_1, e^{i \frac{2\pi}{N_2} k_2} \right) &= \frac{A}{B(k_2) - C(k_2)(z_1 + z_1^{-1})} = \\
 &= \frac{\beta(k_2)}{1 - \alpha(k_2)z_1^{-1}} \times \frac{\beta(k_2)}{1 - \alpha(k_2)z_1} = G^+(z_1)G^-(z_1),
 \end{aligned} \tag{3.39}$$

где $\alpha(k_2)$, $\beta(k_2)$ – коэффициенты, определяемые в процессе факторизации; $G^+(z_1)$ – передаточная функция ЛИС-системы, реализуемой в прямом времени; $G^-(z_1)$ – передаточная функция ЛИС-системы, реализуемой в обратном времени.

Таким образом, получаем следующую спектрально-рекуррентную процедуру фильтрации:

1. Из исходного поля $g(n_1, n_2)$ вычисляется полуспектр по строкам $g^*(n_1, k_2)$ (N_1 раз выполняется N_2 -точечное ДПФ (3.33)).

2. Производится фильтрация в полуспектральной области «в прямом времени»:

$$\hat{f}^*(n_1, k_2) = \alpha(k_2)\hat{f}^*(n_1 - 1, k_2) + \beta(k_2)g^*(n_1, k_2). \tag{3.40}$$

3. То, что получилось, фильтруется «в обратном времени»:

$$\hat{f}^*(n_1, k_2) = \alpha(k_2)\hat{f}^*(n_1 + 1, k_2) + \beta(k_2)\hat{f}^*(n_1, k_2). \tag{3.41}$$

4. Результат переводится в пространственную область (см. (3.34), т.е. опять N_1 раз выполняется N_2 -точечное обратное ДПФ).

Спектрально-рекуррентная реализация оптимального фильтра по числу арифметических операций оказывается проще, чем ре-

лизация с использованием двумерного ДПФ (примерно в два раза). И, что существенно, здесь нет транспонирования матрицы при выполнении ДПФ. К недостатку можно отнести то, что, для простоты расчета фильтра, крайне желательна разделимость всех АКФ и импульсной характеристики искажающей системы (передаточная функция фильтра должна быть дробно-рациональной). Данный алгоритм является асимптотически оптимальным при $N_2 \rightarrow \infty$ (по вертикали матрица может считаться бесконечной).

3.3 Двумерные линейные субоптимальные КИХ-фильтры

Рассмотрим построение субоптимального двумерного линейного КИХ-фильтра путем переноса методики расчета для одномерного случая. Значения импульсной характеристики непосредственно определяются из системы уравнений (3.16), включающей в себя уравнение Винера-Хопфа и ограничения на область ненулевых значений импульсной характеристики. В данном случае область D представляет собой двумерное «окно», содержащее конечное (обычно малое) число отсчетов. Поэтому процедура расчета КИХ-фильтра оказывается достаточно простой: нужно всего лишь решить систему линейных алгебраических уравнений небольшой размерности.

Пример 3.3. Пусть имеет место модель наблюдения (3.26), полезный сигнал имеет биэкспоненциальную изотропную АКФ (3.27), шум – белый, статистически независимый от сигнала, с дисперсией σ_v^2 .

Рассчитаем КИХ-фильтр для окна из пяти точек:

$$D = \{(0,0), (0,1), (1,0), (-1,0), (0,-1)\}. \quad (3.42)$$

Учтем предварительно, что в нашем примере:

$$R_g(k_1, k_2) = R_f(k_1, k_2) + R_v(k_1, k_2), \quad (3.43)$$

$$R_{fg}(k_1, k_2) = R_f(k_1, k_2). \quad (3.44)$$

Учтем также, что, как правило, двумерный сигнал (например, функция яркости изображения) обладает изотропными статистическими свойствами в перпендикулярных направлениях, и поэтому, очевидно, импульсная характеристика искомого КИХ-фильтра будет соответствовать симметричной «маске» (рисунок 3.1) всего с двумя различными числовыми значениями:

$$a = h_{\text{восст}}(0,0)$$

$$b = h_{\text{восст}}(0,1) = h_{\text{восст}}(0,-1) = h_{\text{восст}}(1,0) = h_{\text{восст}}(-1,0). \quad (3.45)$$

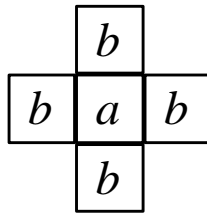


Рис. 3.1. Симметричная маска

Строим систему линейных алгебраических уравнений для определения коэффициентов фильтра. Из уравнения Винера-Хопфа, входящего в (3.16), при $n_1 = n_2 = 0$ получаем:

$$h_{\text{восст}}(0,0)R_g(0,0) + h_{\text{восст}}(0,1)R_g(0,-1) +$$

$$+ h_{\text{восст}}(1,0)R_g(-1,0) + h_{\text{восст}}(0,-1)R_g(0,1) + , \quad (3.46)$$

$$+ h_{\text{восст}}(-1,0)R_g(1,0) = R_{fg}(0,0)$$

или, принимая во внимание новые обозначения:

$$a\left(\sigma_f^2 + \sigma_v^2\right) + 4b\sigma_f^2\rho = \sigma_f^2. \quad (3.47)$$

Из уравнения Винера-Хопфа при $n_1=1, n_2=0$ получаем, что

$$\begin{aligned} h_{\text{восст}}(0,0)R_g(1,0) + h_{\text{восст}}(0,1)R_g(1,-1) + \\ + h_{\text{восст}}(1,0)R_g(0,0) + h_{\text{восст}}(0,-1)R_g(1,1) + \\ + h_{\text{восст}}(-1,0)R_g(2,0) = R_{fg}(-1,0) \end{aligned} \quad (3.48)$$

или в новых обозначениях:

$$a\sigma_f^2\rho + b\left(\sigma_f^2\rho^2 + \left[\sigma_v^2 + \sigma_f^2\right] + \sigma_f^2\rho^2 + \sigma_f^2\rho^2\right) = \sigma_f^2\rho. \quad (3.49)$$

Остальные уравнения будут повторять это. Итак, снова вводя обозначение d^2 для отношения сигнал/шум по мощности, получаем систему:

$$\begin{cases} a\left(\frac{1}{d^2} + 1\right) + 4b\rho = 1; \\ a\rho + b\left(\frac{1}{d^2} + 1 + 3\rho^2\right) = \rho. \end{cases} \quad (3.50)$$

Решая систему, получаем:

$$a = \frac{\left(\frac{1}{d^2} + 1\right) - \rho}{\left(\frac{1}{d^2} + 1\right)\left(\frac{1}{d^2} + 1 + 3\rho^2\right) - 4\rho^2}; \quad (3.51)$$

$$b = \frac{\frac{1}{d^2}\rho}{\left(\frac{1}{d^2} + 1\right)\left(\frac{1}{d^2} + 1 + 3\rho^2\right) - 4\rho^2}. \quad (3.52)$$

Удобство восстановления двумерного сигнала (изображения) КИХ-фильтром заключается в использовании универсальной процедуры линейной обработки изображения скользящим окном

(«маской»). Как уже отмечалось, для того, чтобы эта процедура была достаточно простой, нужно брать маску небольшого размера: 3×3 или 5×5 . При этом для определения оптимальных коэффициентов маски нужно решить систему уравнений соответственно 9-го или 25-го порядка. Если окно симметрично, а изображение обладает изотропными статистическими свойствами, то расчет дополнительно упрощается: при таких условиях для окна 3×3 имеем только 3 различных коэффициента, а для окна 5×5 – шесть. Простота расчета – это тоже достоинство КИХ-фильтра.

На практике измерение или теоретический расчет корреляционных функций не всегда возможны. Поэтому часто используют непараметрический подход к фильтрации. При этом учитывается тот факт, что спектр шума содержит более высокие пространственные частоты, чем спектр идеального изображения. Поэтому любая низкочастотная фильтрация может служить эффективным средством подавления шумов. Приведем типичные примеры сглаживающих масок размером 3×3 :

$$A_1 = \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad A_2 = \frac{1}{10} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad A_3 = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}. \quad (3.53)$$

Коэффициенты данных масок нормированы ($\sum_{k_1, k_2 \in D} \sum h(k_1, k_2) = 1$) с тем, чтобы процедура подавления по-

мех не вызывала смещения средней яркости обработанного изображения относительно исходного.

Маски (3.53) отличаются степенью сглаживания высокочастотных шумов (у маски A_1 она максимальная, у A_3 – минимальная). Выбор маски обычно производится экспериментально. Следует учитывать, что при увеличении степени сглаживания шумов происходит также подавление высокочастотной составляющей

полезного сигнала, что вызывает исчезновение мелких деталей и размазывание контуров изображения.

Если требуемая степень сглаживания шума с применением маски размера 3×3 не достигается, то следует использовать сглаживающие маски больших размеров (5×5 , 7×7 , ...), хотя они используются редко, т.к. для них прямая свертка выполняется достаточно долго. Исключение составляет простое усреднение по окну $K \times K$, которое в случае больших окон может быть реализовано в рекурсивной форме.

Завершая краткий обзор «классических» линейных методов восстановления, следует отметить, что они не полностью решают проблему восстановления изображений. Это связано с несколькими причинами.

Во-первых, как известно из теории оценивания, методы оптимальной линейной фильтрации являются *оптимальными вообще* только для гауссовых сигналов и шумов. Изображения таковыми чаще всего не являются.

Во-вторых, традиционно используемый в таких методах среднеквадратичный критерий качества восстановления плохо согласуется со свойствами зрения, а также со многими целевыми функциями обработки изображений.

В-третьих, не всякие искажения описываются введенной выше линейной моделью наблюдения.

И, в-четвертых, не всегда известны те характеристики сигналов и искажений, которые нужны для построения фильтра.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Теоретические основы цифровой обработки сигналов и изображений. Детерминированные модели: учебное пособие / А.Ю. Денисова, В.В. Сергеев, М.А. Чичева, Р.Р. Юзькив. – Самара: Изд-во Самарского университета, 2023. – 168 с.
2. Вентцель, Е.С. Теория вероятностей / Е.С. Вентцель – Юстиция, 2018. – С. 658-658.
3. Дауни, А.Б. Цифровая обработка сигналов на языке Python / А.Б. Дауни. – Москва: ДМК Пресс. – 2017. – 160 с.
4. Землянухин, П.А. Теория сигналов: учебное пособие / П.А. Землянухин. – Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2022. – 136 с.
5. Лайонс, Р. Цифровая обработка сигналов / Р. Лайонс. – Москва: Бином, 2015. – 656 с.
6. Магазинникова, А.Л. Основы цифровой обработки сигналов: учебное пособие / А.Л. Магазинникова. – Санкт-Петербург: Лань, 2021. – 132 с.
7. Оппенгейм, А. Цифровая обработка сигналов / А. Оппенгейм, Р. Шафер. – Москва: Техносфера, 2012. – 1048 с.
8. Сергиенко, А.Б. Цифровая обработка сигналов / А.Б. Сергиенко. – Санкт-Петербург: БХВ-Петербург, 2011.
9. Смит, С. Цифровая обработка сигналов. Практическое руководство для инженеров и научных работников / С. Смит. – Москва: ДМК Пресс, 2018. – 718 с.
10. Умняшкин, С.В. Основы теории цифровой обработки сигналов: учебное пособие / С.В. Умняшкин. – Москва: Техносфера, 2019. – 550 с.

Учебное издание

*Сергеев Владислав Викторович,
Чичева Марина Александровна,
Юзькив Руслан Романович,
Денисова Анна Юрьевна*

**ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЦИФРОВОЙ
ОБРАБОТКИ СИГНАЛОВ И ИЗОБРАЖЕНИЙ.
ВЕРОЯТНОСТНЫЕ МОДЕЛИ**

Учебное пособие

Редакционно-издательская обработка
издательства Самарского университета

Подписано в печать 15.05.2024. Формат 60x84 1/16.

Бумага офсетная. Печ. л. 5,25.

Тираж 120 экз. (1-й з-д 1-27). Заказ № .

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)
443086, Самара, Московское шоссе, 34.

Издательство Самарского университета.
443086, Самара, Московское шоссе, 34.