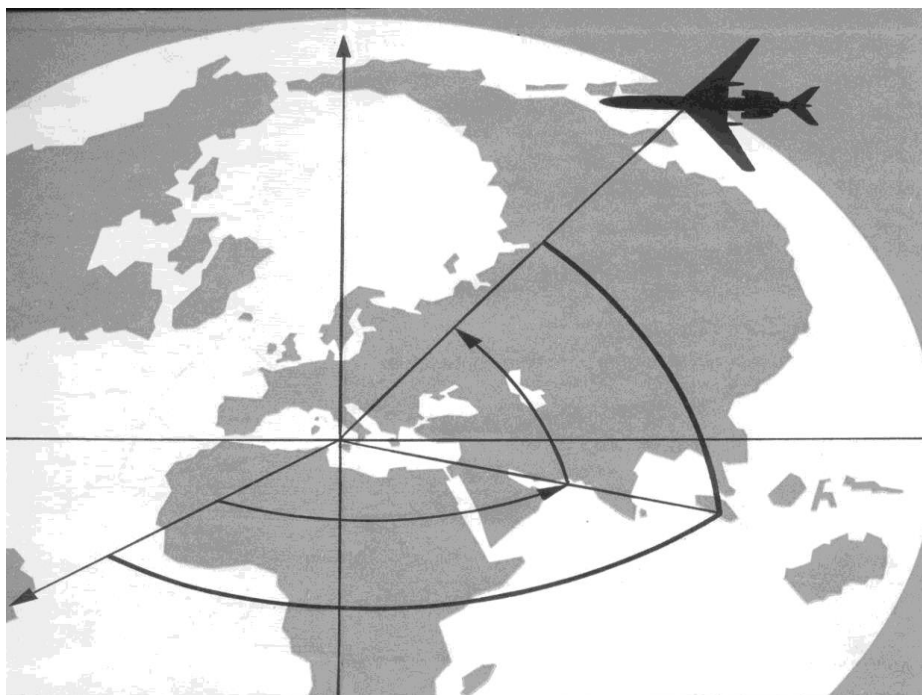


Кичак В. М., Воловик А. Ю., Шутило М. А., Червак О. П.

РАДІОТЕХНІЧНІ СИСТЕМИ

(ОСНОВИ ПРОЕКТУВАННЯ. ЧАСТИНА 2)



Міністерство освіти і науки України
Вінницький національний технічний університет

Кичак В. М., Воловик А. Ю., Шутило М. А., Червак О. П.

**РАДІОТЕХНІЧНІ СИСТЕМИ
(ОСНОВИ ПРОЕКТУВАННЯ. ЧАСТИНА 2)**

Навчальний посібник

Вінниця
ВНТУ
2018

УДК 621.396(075)
Р15

Автори:

Кичак В. М., Воловик А. Ю, Шутило М. А., Червак О. П.

Рекомендовано до друку Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України (протокол № ____ від _____ р.)

Рецензенти:

В. С. Осадчук, доктор технічних наук, професор

С. М. Злепко, доктор технічних наук, професор

О. Б. Шарпан, доктор технічних наук, професор

Р15 **Радіотехнічні системи. Основи проектування. Частина 2 :**
навчальний посібник / В. М. Кичак, А. Ю. Воловик, М. А. Шутило,
О. П. Червак – Вінниця : ВНТУ, 2018. – 99 с.

В посібнику розглянуті питання аналізу та синтезу радіотехнічних пристроїв, систем та комплексів різного цільового призначення з використанням сучасних інформаційних технологій, визначені типові етапи проектування.

Призначений для студентів спеціальності «Радіотехніка» всіх кваліфікаційних рівнів.

УДК 621.396(075)

ЗМІСТ

ЧАСТИНА ПЕРША

ОСНОВИ КЛАСИЧНОЇ ТЕОРІЇ РАДІОТЕХНІЧНИХ СИСТЕМ

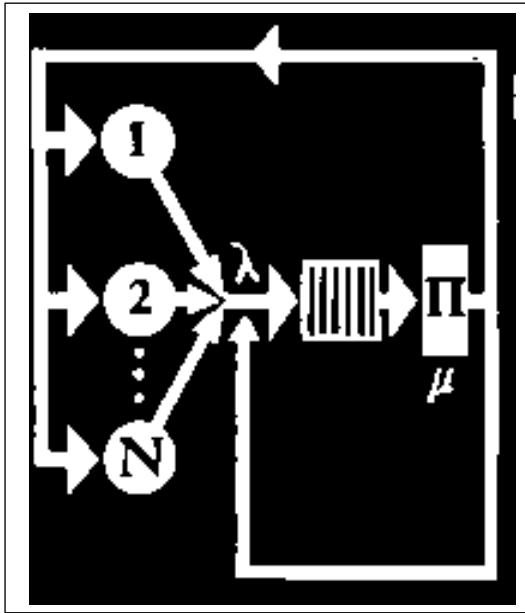
1 МАТЕМАТИЧНІ МЕТОДИ АНАЛІЗУ ПРОЦЕСІВ У НЕЛІНІЙНИХ БЕЗІНЕРЦІЙНИХ СИСТЕМАХ.....	6
1.1 Методи аналізу детермінованих процесів	6
1.2 Нелінійні безінерційні перетворення випадкових процесів	10
1.3 Викиди випадкових процесів	16
2 ВИПАДКОВІ ПРОЦЕСИ У НЕЛІНІЙНИХ ІНЕРЦІЙНИХ СИСТЕМАХ	19
2.1 Метод лінеаризації	19
2.2 Метод статистичної лінеаризації.....	20
2.3 Застосування рівняння Фоккера-Планка-Колмогорова	23
2.4 Ймовірнісне моделювання радіотехнічних систем на електронних обчислювальних машинах	28
3 МАТЕМАТИЧНИЙ СИНТЕЗ ОПТИМАЛЬНИХ РАДІОТЕХНІЧНИХ СИСТЕМ	30
3.1 Базові положення теорії прийняття статистичних рішень	30
3.2 Задачі виявлення та розпізнавання сигналів	32
3.3 Статистичні оцінки параметрів сигналів	34
3.4 Розрізнення сигналів та синтез сигналів оптимальної форми.....	40
3.5 Лінійна оптимальна фільтрація	44
3.6 Нелінійна оптимальна фільтрація при гаусовій апроксимації	47
3.7 Нелінійна оптимальна фільтрація марковських процесів	52

ЧАСТИНА ДРУГА

СУЧАСНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В РАДІОТЕХНІЧНИХ СИСТЕМАХ

4 БАЗОВІ ПОНЯТТЯ СУЧАСНОЇ СИСТЕМОТЕХНІКИ.....	57
4.1 Основні засади загальної теорії систем та системного підходу	57
4.2 Методи загальної теорії систем	64
5 ЙМОВІРНІСНІ МОДЕЛІ ВИМІРЮВАНИХ ПАРАМЕТРІВ.....	66
5.1 Вступ.....	66
5.2 Математичні моделі випадкових сигналів та процесів	67
6 ОПТИМАЛЬНЕ ОЦІНЮВАННЯ ПАРАМЕТРІВ	77

7 НЕЛІНІЙНА ОПТИМАЛЬНА ФІЛЬТРАЦІЯ	88
7.1 Постанова задачі.....	88
7.2 Нелінійне оцінювання за методом максимуму апостеріорної ймовірності.....	88
ПІСЛЯМОВА	97
ПЕРЕЛІК ПОСИЛАНЬ.....	98



1

ОСНОВИ КЛАСИЧНОЇ ТЕОРІЇ РАДІОТЕХНІЧНИХ СИСТЕМ

- Математичні методи аналізу процесів у нелінійних безінерційних системах
- Випадкові процеси в нелінійних інерційних системах
- Математичний синтез оптимальних радіотехнічних систем

1 МАТЕМАТИЧНІ МЕТОДИ АНАЛІЗУ ПРОЦЕСІВ У НЕЛІНІЙНИХ БЕЗІНЕРЦІЙНИХ СИСТЕМАХ

1.1 Методи аналізу детермінованих процесів

Аналіз нелінійних систем становить великі труднощі навіть для детермінованих зовнішніх діянь, і дотепер не існує єдиного математичного методу, придатного для дослідження будь-яких нелінійних систем при довільних режимах їх роботи. Кожний запропонований метод виявляється досить ефективним, як правило, лише для одного або декількох режимів роботи того чи іншого класу нелінійних систем. Виявляється, що у нелінійній системі, у порівнянні з лінійною, можливий додатковий режим – режим автоколивань. Тому при аналізі можна досліджувати стійкість станів спокою – режим рівноваги або режим усталених автоколивань, перехідні процеси при встановленні або зникненні автоколивань, вплив зовнішньої дії у режимі відсутності або наявності автоколивань.

Ще у більшій мірі різноманітні радіотехнічні нелінійні системи. Сюди відносять усілякі автогенератори коливань синусоїдальної та спеціальної форми, помножувачі та подільники частоти, обмежники, випрямлячі, модулятори і демодулятори, електронні та електромеханічні реле і багато інших нелінійних елементів, пристроїв та систем у цілому. Окрім того, вплив нелінійних характеристик може бути істотним у таких елементах, пристроях і системах, які у нормальному режимі роботи вважалися лінійними. Порядок n диференціального рівняння, яке з достатньою повнотою описує нелінійну систему, також може бути суттєво різним – починаючи від $n = 1$ і закінчуючи $n = 10\text{--}30$ і більше. Відповідно до різноманіття типів нелінійних систем та режимів їх роботи, на сьогоднішній день, відомо щонайменше декілька десятків різних методів досліджень. У даному посібнику розглядається група найпоширеніших методів, які застосовуються до порівняно широких класів нелінійних систем, а саме:

- метод лінеаризації;
- метод гармонічної лінеаризації;
- метод малого параметра;
- метод осереднення;
- метод фазової площини;
- метод інтегральної апроксимації;
- метод математичного моделювання.

У зв'язку з тим, що більшість цих методів уже знайома читачеві з дисциплін, що читалися на рівні бакалаврату, то у даному навчальному посібнику автор вважає доцільним обмежитись коротким оглядом вище перерахованих методів з посиланнями на відповідну літературу.

Метод лінеаризації [6]. Цей метод заснований на тому, що при невеликих збуреннях систему можна вважати умовно лінійною відносно цих збурень і аналізувати її методами, розробленими для лінійних систем. Однак такий метод неприйнятний при великих збуреннях, а у деяких спеціальних випадках (наприклад, при нелінійних характеристиках типу зони нечутливості) і при малих збуреннях. Характерним прикладом застосування методу лінеаризації є дослідження стійкості стаціонарної системи «у малому», тобто при малих відхиленнях від стану рівноваги або періодичного колювання (автоколювання). При застосуванні цього методу, отримують систему з n лінійних диференціальних рівнянь першого порядку (або одного диференціального рівняння n -го порядку). Якщо коефіцієнти рівняння постійні, то питання про стійкість або нестійкість системи розв'язується шляхом застосування одного з методів аналізу стійкості лінійних стаціонарних систем (методу Рауса–Гурвіца, Найквіста та ін.). Іншим прикладом застосування методу лінеаризації є дослідження точності наведення на ціль керованого снаряду при малих рівнях завад та інших чинниках [6].

Метод гармонічної лінеаризації [16]. Цей метод базується на заміні нелінійного елемента лінійним елементом, еквівалентним за дією при гармонійному вхідному сигналі, тобто мають на увазі, що коефіцієнт передачі за першою гармонікою описується виразом

$$S_{cp} = \frac{\dot{I}_1}{U} = S_{cp} e^{j\varphi_{cp}},$$

де \dot{I}_1 – комплексна амплітуда першої гармоніки колювання на виході нелінійного елемента, а U – амплітуда синусоїдального колювання на його вході. Якщо U і \dot{I}_1 мають, відповідно, розмірності напруги і струму, то величину S_{cp} називають середньою крутизною характеристики нелінійного елемента. Очевидно, що такий метод можна застосовувати лише у тому випадку, коли колювання у системі близькі до гармонійних, тобто дією вищих гармонічних складових, що виникають на виході нелінійного елемента, можна знехтувати унаслідок значного їх ослаблення наступними вузькосмуговими фільтрами. Оскільки при заміні нелінійного елемента його лінійним еквівалентом система, у цілому, стає лінійною, то її можна досліджувати кожним з методів, розроблених для лінійних систем. Але за рахунок того, що у розглянутому випадку колювання близькі до синусоїдальних, можливе додаткове спрощення аналізу лінеаризованої системи шляхом застосування методу повільно-мінливих комплексних обвідних. Це еквівалентно укороченню диференціальних або операторних рівнянь. У такому випадку метод гармонічної лінеаризації можна розглядати як реалізацію методу повільно-мінливих комплексних обвідних стосовно нелінійних систем. Вперше метод гармонічної лінеаризації з'явився багато років тому стосовно аналізу лампового генератора або регенератора і отримав назву квазілінійного методу. Пізніше він був розповсюджений на системи

автоматичного регулювання і одержав найменування методу Гольдфарба (методу гармонійного балансу).

Метод малого параметра [17]. Основна ідея методу малого параметра полягає у наступному. Припустимо, що диференціальне рівняння, яке описує поведінку системи, допускає таку форму опису, у якій права частина диференціального рівняння містить у вигляді множника деякий малий параметр ε такий, що при $\varepsilon \rightarrow 0$ розв'язок відомий і дорівнює $x_0(t)$. Тоді при $\varepsilon \neq 0$ розв'язок $x(t)$ диференціального рівняння відшукується у вигляді ряду за степенями цього малого параметра, тобто у вигляді

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \dots \quad (1.1)$$

Якщо обмежитися доданками зі степенем ε не вище першого, то отриманий розв'язок називається першим наближенням. При обліку доданка з ε^2 одержують друге наближення і т.д. Якщо підставити вираз (1.1) у вищезгадане диференціальне рівняння та виконати низку перетворень, то можна отримати рівняння для визначення поправок першого наближення $\varepsilon x_1(t)$, другого наближення $\varepsilon^2 x_2(t)$ і наближень більш високого порядку. Звичайно, друге наближення обчислити значно складніше, ніж перше; третє наближення ще суужніше і т.д. Однак у більшості випадків достатня точність досягається уже в результаті обчислення першого наближення. Нульове наближення $x_0(t)$, що має місце при $\varepsilon=0$ називають номінальним розв'язком. Наприклад, у випадку системи другого порядку, її диференціальне рівняння часто вдається записати у вигляді

$$\frac{d^2 x}{dt^2} + \omega_0^2 x = \varepsilon f\left(x, \frac{dx}{dt}\right), \quad (1.2)$$

а нульове наближення, що є розв'язком рівняння

$$\frac{d^2 x}{dt^2} + \omega_0^2 x = 0,$$

має вигляд $x_0(t) = A \sin \omega_0 t$. Це і є номінальний розв'язок. Існують численні різновиди методу малого параметра: метод Пуанкаре, метод Ляпунова, метод Ван-Дер Поля, метод Н. М. Крилова, І. Н. Боголюбова та ін., які різняться вибором типу твірної системи та способами знаходження поправок першого наближення, другого наближення і т.д. Швидко і правильно вибрати твірне рівняння, тобто визначити, який саме параметр системи слід вибрати у якості малого параметра, як правило, вдається тільки для систем другого порядку, а в системах вищих порядків методи малого параметра значною мірою втрачають простоту і фізичну наочність. За досить загальних умов перше наближення $x_0(t) + \varepsilon x_1(t)$ збігається з розв'язком,

отриманим методом лінеаризації, а при врахуванні поправок другого наближення можна одержати більш точний розв'язок. Отже, методи малого параметра значно складніші методів лінеаризації, але дозволяють у низці випадків отримувати більш точні розв'язки і, зокрема, визначати межі застосування методу гармонічної лінеаризації.

Метод осереднення [1,16]. При застосуванні методів малого параметра часто виникають істотні труднощі внаслідок того, що права частина диференціального рівняння (або праві частини системи диференціальних рівнянь), у яку в якості множника входить малий параметр ε , явно залежить від часу t . Наприклад,

$$\frac{d^2x}{dt^2} + \omega_0^2 x = \varepsilon f\left(x, \frac{dx}{dt}, t\right).$$

У цьому разі, з метою спрощення розв'язку задачі, застосовують метод осереднення, який заснований на тому, що права частина диференціального рівняння замінюється її середнім значенням, причому осереднення проводиться лише за змінною часу t . Фундаментальне обґрунтування методу осереднення та умов його застосування детально розглядалось у [16].

Метод фазової площини [1,6]. Якщо процеси, що протікають у системі, повністю описуються завданням двох змінних x і y , то вони можуть бути відображені у декартовій координатній системі з осями x і y . При цьому площина (x, y) називається фазовою площиною, а кожна крапка з поточними координатами (x, y) — зображувальною крапкою. Так як x і y є функціями часу, то з часом зображувальна крапка описує у фазовій площині криву, яку прийнято називати фазовою траєкторією. Сукупність фазових траєкторій, що відповідають різним значенням початкових умов, називають фазовим портретом системи. Метод фазової площини знаходить широке застосування при дослідженні систем другого порядку. При цьому, якщо рівняння системи презентовано у вигляді

$$\frac{d^2x}{dt^2} = F\left(x, \frac{dx}{dt}\right),$$

то на осях координат відкладають змінні x і $y = \frac{dx}{dt}$. Дослідження поведінки фазових траєкторій дозволяє відповісти на низку важливих питань, розв'язок яких не може бути знайдений описаними вище наближеними методами, зокрема, виявити можливі види коливальних у даній системі при різних значеннях початкових умов, а також оцінити стійкість станів рівноваги та періодичного руху при наявності великих збурень. Найбільш повний виклад методів дослідження фазових портретів систем другого порядку подано [16]. Якщо система описується не двома, а трьома змінними (система третього порядку), то її поведінку можна розглядати у фазовому просторі (x, y, z) . При цьому фазовий портрет виявляється досить складним для по-

будови та дослідження. У зв'язку з цим, метод фазового простору отримав значно менший розвиток і менш поширений.

Методи інтегральної апроксимації [1, 9, 16]. Ці методи засновані на такій апроксимації статичної вольт-амперної характеристики нелінійного елемента або ділянок цієї характеристики, за якої виявляється можливим знаходження точного аналітичного розв'язку диференціального рівняння, що описує поведінку системи. При цьому, якщо характеристику довелося розбивати на декілька ділянок, то проводиться «зшивання» отриманих розв'язків шляхом прирівнювання знайдених розв'язків і їх похідних (якщо вони існують) на границях сусідніх ділянок. Найпоширенішою є кусково-лінійна апроксимація, де кожній ділянці характеристики відповідає певне лінійне диференціальне рівняння, розв'язок якого знаходиться найбільш просто. Однак при збільшенні числа ділянок трудомісткість обчислень різко зростає. Тому даний метод застосовують звичайно лише в тих випадках, коли характеристика може бути розбита на невелике число ділянок, зокрема при дослідженні систем з релейними елементами,

Математичне моделювання. У багатьох випадках зробити досить детальне теоретичне дослідження нелінійної системи без допомоги обчислювальних машин не вдається і доводиться застосовувати математичне моделювання — приводити рівняння, що описують систему, до вигляду зручному для розв'язку на цифровій або аналоговій обчислювальній машині, і робити необхідні дослідження за допомогою цієї машини. Більш докладно ці питання розглядаються в [5] стосовно до випадку, коли вхідні діяння можуть містити не тільки детерміновані, але випадкові складові.

1.2 Нелінійні безінерційні перетворення випадкових процесів

Розглянемо загальні методи визначення ймовірнісних характеристик випадкових процесів після нелінійного безінерційного перетворення. Оскільки значення процесу на виході нелінійного елемента $\zeta(t)$ визначаються тільки значенням процесу на вході $\xi(t)$ у той же момент часу

$$\zeta(t) = f[\xi(t)],$$

то для визначення функції розподілу випадкового процесу на виході нелінійного елемента $W_n(y_1, \dots, y_n, t_1, \dots, t_n)$ досить мати функцію розподілу вхідного процесу $W_n(x_1, \dots, x_n, t_1, \dots, t_n)$. Далі слід використовувати відомі у теорії ймовірностей загальні формули заміни змінних у функціях розподілу та обчислення матриці перетворення Якобі [11]. Іноді ці обчислення можна спростити, якщо спочатку знаходити не самі функції розподілу, а їх характеристичні функції. Слід зауважити, що визначення функцій розподілу процесу $\zeta(t)$ є порівняно простим і зводиться до заміни змінних у $W_n(x_1, \dots, x_n, t_1, \dots, t_n)$ та підрахунку якобіана перетворення тільки у тому випадку, коли нелінійне перетворення $\zeta(t) = f[\xi(t)]$ є взаємно однозначним. Однак на практиці часто трапляється так, що обернена функція $f[\xi(t)]$

багатозначна або навіть нескінченнозначна, а іноді й зовсім не існує. Тоді визначення багатомірних функцій розподілу стає досить громіздким, а в деяких випадках вимагає застосування спеціальних математичних методів [17]. Після того, як знайдено, щонайменше, двовимірний розподіл процесу на виході нелінійного елемента $W_2(y_1, y_2, t_1, t_2)$ функцію моментів цього процесу можна, з принципової точки зору, визначати шляхом обчислення подвійного інтегралу

$$B_\zeta(t_1 t_2) = \iint_{-\infty}^{\infty} y_1 y_2 W_2(y_1, y_2, t_1, t_2) dy_1 dy_2.$$

Однак, якщо цікавляться тільки функціями моментів (а не розподілом ймовірностей) після нелінійного безінерційного перетворення, то простіше використовувати співвідношення, які встановлюють безпосередню залежність моментної функції процесу $\zeta(t)$ від двовимірної функції розподілу процесу $\xi(t)$ та характеристики нелінійності $f[\xi(t)]$

$$B_\zeta(t_1 t_2) = \iint_{-\infty}^{\infty} f(x_1) f(x_2) W_2(x_1, x_2, t_1, t_2) dx_1 dx_2. \quad (1.3)$$

Існує два загальні методи обчислення цього інтегралу: так званий, прямий метод і метод характеристичних функцій, а для нормальних випадкових процесів — також метод похідних.

Прямий метод. Цей метод заснований на ортогональній розвиненні у ряд двовимірної густини розподілу ймовірностей $W_2(x_1, x_2, t_1, t_2)$ процесу на вході нелінійного елемента. Припустимо, що $W_1(x, t)$ — одновимірна функція розподілу, яка відповідає типу двовимірної густини розподілу ймовірностей $W_2(x_1, x_2, t_1, t_2)$ процесу на вході нелінійного елемента. Приймемо $W_1(x, t)$ у якості вагової функції і збудуємо сукупність нормованих ортогональних поліномів $Q_n(x, t)$, які повинні задовольняти умовам ортогональності

$$\int_{-\infty}^{\infty} W_1(x, t) Q_n(x, t) Q_m(x, t) dx = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases}$$

Тоді двовимірну функцію розподілу $W_2(x_1, \dots, x_2, t_1, t_2)$ можна розвинути у подвійний ряд за цими ортогональними поліномами

$$\begin{aligned} W_2(x_1, x_2, t_1, t_2) &= W_1(x_1, t_1) W_1(x_2, t_2) \times \\ &\times \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn}(t_1, t_2) Q_m(x_1, t_1) Q_n(x_2, t_2). \end{aligned} \quad (1.4)$$

Коефіцієнти $a_{mn}(t_1, t_2)$ можуть бути визначені шляхом перемноження обох частин (1.4) на $Q_k(x_1, t_1)Q_l(x_2, t_2)$ та наступним інтегруванням з використанням умови ортогональності. Тоді

$$a_{mn}(t_1, t_2) = \iint_{-\infty}^{\infty} W_2(x_1, x_2, t_1, t_2) Q_m(x_1, t_1) Q_n(x_2, t_2) dx_1 dx_2.$$

У багатьох, практично важливих, випадках при $m \neq n$ $a_{mn} = 0$. Для цього класу функцій розподілу вищенаведені формули спрощуються

$$\begin{aligned} W_2(x_1, x_2, t_1, t_2) &= W_1(x_1, t_1) W_1(x_2, t_2) \times \\ &\times \sum_{x_0}^{\infty} a_n(t_1, t_2) Q_n(x_1, t_1) Q_n(x_2, t_2). \\ a_n(t_1, t_2) &= \iint_{-\infty}^{\infty} W_2(x_1, x_2, t_1, t_2) Q_n(x_1, t_1) Q_n(x_2, t_2) dx_1 dx_2. \end{aligned}$$

Наступна підстановка дозволяє розділити змінні інтегрування та приводить до остаточної формули

$$B_{\zeta}(t_1, t_2) = \sum_{n=0}^{\infty} C_n(t_1) C_n(t_2) a_n(t_1, t_2);$$

де

$$C_n(t) = \int_{-\infty}^{\infty} f(x) Q_n(x, t) W_1(x, t) dx.$$

Важливо підкреслити, що коефіцієнти $C_n(t)$ залежать тільки від типу характеристики нелінійності та одновимірного розподілу процесу на вході нелінійного елемента, у той час, як функції $a_n(t_1, t_2)$ не залежать від виду нелінійності і визначаються тільки двовимірним розподілом вхідного процесу. У практичних розрахунках функції $a_n(t_1, t_2)$ бувають заздалегідь відомими і необхідно лише визначити $C_n(t)$, відповідно до конкретного типу нелінійності. Наприклад, для нормального випадкового процесу, що являє собою адитивну суміш детермінованого процесу і нормального, стаціонарного шуму [11]

$$a_n(t_1, t_2) = \frac{1}{n!} R^n(t_2 - t_1)$$

де $\sigma^2 R(\tau)$ — кореляційна функція шуму. У цьому випадку сукупність ортогональних поліномів $Q_n(x, t)$ являє собою поліноми Ерміта.

Метод характеристичних функцій (контурних інтегралів). Цей метод заснований на тому, що характеристики деяких нелінійних елементів можна описувати за допомогою контурного інтегралу

$$f(x) = \frac{1}{2\pi} \int_c g(ju) e^{jxu} du, \quad (1.5)$$

де

$$g(ju) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-jxu} dx,$$

Якщо підставити (1.5) в (1.3) та змінити порядок інтегрування, то приходимо до формули

$$B_{\zeta}(t_1 t_2) = \frac{1}{4\pi^2} \iint_{c_1 c_2} g(ju_1) g(ju_2) \theta_2(u_1, u_2, t_1, t_2) du_1 du_2,$$

де $\theta_2(\theta, u_2, t_1, t_2)$ — двовимірна характеристична функція процесу на вході нелінійного елементу. Розділення змінних у цьому виразі виконується шляхом розвинення функції $\theta_2(u_1, u_2, t_1, t_2)$ у ряд. Не існує загальних рекомендацій про те, у яких випадках доцільніше використовувати саме прямий метод обчислення функції моментів, а у яких — метод контурних інтегралів. Обчислення коефіцієнтів розвинення у ряд залежно від статистичних властивостей процесу на вході нелінійного елементу або від типу нелінійності може виявитися простіше або одним, або іншим методом. Іноді використання методу контурних інтегралів для нестационарних процесів може виявитися більш привабливим, якщо надалі буде необхідне додаткове осереднення за часом для функції моментів вихідного процесу.

Метод похідних. Коли на вході нелінійного безінерційного елементу діє нормальний випадковий процес, є можливість дуже просто обчислити не функцію моментів процесу на виході нелінійного елементу, а похідні від цієї функції по функції кореляції вхідного нормального процесу, після чого шукана функція моментів знаходиться шляхом елементарного інтегрування. Цей спосіб визначення функцій моментів після нелінійного перетворення, запропонований Прайсом [6], і є наслідком застосування методу характеристичних функцій до нормальних випадкових процесів. У зв'язку з важливістю ролі нормальних випадкових процесів у різних прикладних дослідженнях, цей метод розглянемо більш докладніше. Безпосереднє диференціювання кратного інтегралу від характеристичної функції нормального процесу $\theta_2(u_1, u_2, t_1, t_2)$ по параметру R приводить до співвідношення

$$\begin{aligned} \frac{d^k B_{\xi}(\tau)}{dR^k(\tau)} &= \frac{\sigma^{2k-2}}{2\pi\sqrt{(1-R^2(\tau))}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^{(k)}(x_1) f^{(k)}(x_2) \times \\ &\times \exp\left\{ \frac{1}{2\sigma^2(1-R^2(\tau))} [(x_1 - m_1)^2 - 2R(\tau)(x_1 - m_1)(x_2 - m_1) + (x_2 - m_1)^2] \right\} dx_1 dx_2 \end{aligned} \quad (1.6)$$

Формула (1.6) лежить у основі розглянутого методу похідних. Якщо функція $f(x)$ складена із фрагментів поліномів, то при деякому значенні k по-

хідна $f^{(k)}(x)$ буде дорівнює сумі дельта-функцій і обчислення інтегралу (1.6) стає елементарним, якщо скористатися фільтруючою властивістю дельта-функції та її похідних [1]. У якості ілюстрації розглянемо один, окремий вид нелінійного перетворення стаціонарного випадкового процесу, що являє собою особливий інтерес для радіотехнічних систем, які використовують для передачі інформації електромагнітні коливання високої частоти. Мова йде про представлення вузько смугових процесів у вигляді

$$\xi(t) = E(t) \cos[\omega_0 t - \varphi(t)],$$

де $E(t)$, $\varphi(t)$ – обвідна і фаза процесу $\xi(t)$, причому моментні функції обвідної та фази повільно змінюються порівняно з $\cos[\omega_0 t]$ (тобто, їх енергетичні спектри зосереджені у низькочастотній області). Функціональний зв'язок обвідної та фази з вхідним процесом задається співвідношеннями

$$E(t) = \sqrt{\xi^2(t) + \eta^2(t)},$$

$$\varphi(t) = \left[\omega_0 t - \operatorname{arctg} \frac{\eta(t)}{\xi(t)} \right],$$

де $\eta(t)$ — так званий, спряжений з $\xi(t)$ процес, що є перетворенням Гільберта від $\xi(t)$

$$\eta(t) = \frac{1}{\pi} \lim_{T \rightarrow \infty} \int_{-T}^T \frac{\xi(\tau)}{t - \tau} d\tau.$$

Як видно з наведених формул, залежності обвідної і фази від першопочаткового процесу достатньо складні. Однак, оскільки вхідний процес є вузько смуговим, то використовуючи цю властивість, його можна представляти у вигляді суми двох квадратурних (ортогональних) складових

$$\xi(t) = A(t) \cos \omega_0 t + C(t) \sin \omega_0 t$$

де $A(t)$ і $C(t)$ випадкові процеси, енергетичні спектри яких однакові і їх можна отримати шляхом зсуву вузько смугового енергетичного спектру $S(\omega)$ у область нижніх частот на величину ω_0 . Відповідно до теореми Хінчина–Вінера, моментна функція кожної із квадратурних складових отримується косинус-перетворенням Фур'є від $S(\omega - \omega_0)$, а взаємна моментна функція квадратурних складових — синус-перетворенням Фур'є від $S(\omega - \omega_0)$. Процес, спряжений з $\xi(t)$ також може бути вираженим через квадратурні складові

$$\eta(t) = A(t) \cos \omega_0 t - C(t) \sin \omega_0 t.$$

З цього випливає, що

$$E(t) = \sqrt{A^2(t) + C^2(t)},$$

$$\varphi(t) = \operatorname{arctg} \frac{C(t)}{A(t)}.$$

Якщо $\xi(t)$ – нормальний вузько смуговий стаціонарний випадковий процес, то випадкові процеси $A(t)$, $C(t)$ також нормальні. Використання вищеведених формул, дозволяє отримати функції розподілу обвідної та фази випадкового процесу. Нехай $W_{2n}(x_1, \dots, x_n, y_1, \dots, y_n, t_1, \dots, t_n)$ – спільний розподіл значень $A(t)$, $C(t)$ у n моментах часу. Для того щоб знайти функції розподілу фази і обвідної перейдемо у зазначеному спільному розподілі до полярної системи координат

$$x_k = r_k \cos \eta_k, \quad y_k = r_k \sin \eta_k, \quad k = 1, \dots, n.$$

Після такої заміни замість спільної функції розподілу від змінних $x_1, \dots, x_n, y_1, \dots, y_n$ отримуємо $2n$ -вимірну спільну функцію розподілу обвідної і фази залежно від нових змінних $r_1, \dots, r_n, \eta_1, \dots, \eta_n$

$$\begin{aligned} & W_{2n}(r_1, \dots, r_n, \eta_1, \dots, \eta_n, t_1, \dots, t_n) = \\ & = D_n \times W_{2n}(r_1 \cos \eta_1, \dots, r_n \cos \eta_n, r_1 \sin \eta_1, \dots, r_n \sin \eta_n, t_1, \dots, t_n), \end{aligned} \quad (1.7)$$

де D_n — якобіан перетворення, рівний $D_n = \sum_{i=1}^n r_i$. Функцію розподілу обвідної можна отримати шляхом n -кратного інтегрування цього розподілу за змінним η_1, \dots, η_n

$$\begin{aligned} & W_{2n}(r_1, \dots, r_n, t_1, \dots, t_n) = \\ & = \int_0^{2\pi} \dots \int_0^{2\pi} W_{2n}(r_1, \dots, r_n, \eta_1, \dots, \eta_n, t_1, \dots, t_n) d\eta_1, \dots, d\eta_n, \quad r_i > 0. \end{aligned}$$

Функція розподілу фаз випадкового процесу отримується із (1.7) шляхом інтегруванням по змінним r_1, \dots, r_n

$$\begin{aligned} & W_{2n}(\eta_1, \dots, \eta_n, t_1, \dots, t_n) = \\ & = \int_0^{2\pi} \dots \int_0^{2\pi} W_{2n}(r_1, \dots, r_n, \eta_1, \dots, \eta_n, t_1, \dots, t_n) dr_1, \dots, dr_n, \quad |\eta_i| \leq \pi. \end{aligned}$$

Слід зауважити, що через дуже складну нелінійну функціональну залежність обвідної $E(t)$ та фази $\varphi(t)$ від $\xi(t)$ використання формули (1.3) для визначення моментних функцій стає не можливим. Тому для визначення моментних функцій фази та обвідної, спочатку доцільно по вищезазначеним формулам визначити двовимірні густини розподілу ймовірностей, а потім обчислити відповідні моментні функції $B_\zeta(t_1 t_2)$, використовуючи загальні співвідношення (1.3). Так, наприклад, моментна функція обвідної нормального стаціонарного вузько смугового випадкового процесу представляється рядом

$$B_E(\tau) = \sigma^2 \sum_{n=0}^{\infty} C_n^2 R_0^{2n}(\tau),$$

$$C_n = \int_0^{\infty} x^2 L_n^{(0)}(x^2/2) e^{-x^2/2} dx,$$

$$R_0^2(\tau) = R_c^2(\tau) + R_s^2(\tau),$$

де σ^2 – дисперсія вихідного процесу, $\sigma^2 R_c^2(\tau)$ і $\sigma^2 R_s^2(\tau)$ – функції кореляції та взаємної кореляції складових $A(t)$ і $C(t)$, $L_n^{(0)}(y)$ — поліноми Лагерра.

Викладена вище методика визначення ймовірнісних характеристик фази та обвідної вузько смугових випадкових процесів, також дозволяє вирішувати, і більш складні завдання. Наприклад, визначати розподіл ймовірностей похідних від обвідної $E'(t)$ або миттєвої частоти $\varphi(t)$. Ці задачі розглядаються у роботі [1,16].

1.3 Викиди випадкових процесів

Для розв'язку дуже багатьох важливих практичних завдань необхідно знати функцію розподілу або числові характеристики тривалості викидів випадкового процесу $\xi(t)$. Під *тривалістю викиду* $\xi_{\text{в}}$ розуміють відрізок часу, протягом якого $\xi(t)$ перевищує заданий граничний рівень $x=x_0$. Поряд із цим, становить інтерес тривалість інтервалу між викидами $\xi_{\text{и}}$, тобто відрізка часу, протягом якого $\xi(t)$ не перевищує заданий рівень $x=x_0$. Іноді замість розподілу тривалості викидів вивчають розподіл точок перетинання цієї функції із заданою прямою. У більш загальному випадку, може розглядатися задача про визначення розподілу точок перетинання $\xi(t)$ із заданою функцією $f(t)$ або про тривалість викидів над заданою функцією $f(t)$. Подальшим розвитком таких завдань є задача про розподіл часу перебування випадкових процесів між двома рівнями (або функціями часу). Певний інтерес становить задача про розподіл ймовірностей тривалості інтервалу між фіксованим моментом часу, у якому випадковий процес приймає задане значення, і моментом першого перетинання цим процесом рівня $x=x_0$ з певним знаком похідної (наприклад, перетинанням рівня знизу нагору). Через актуальність проблеми, задача знаходження ймовірнісних характеристик викидів випадкових процесів тривалий час привертала та ще й досі привертає увагу математиків і інженерів. Райс [17] усередині минулого століття встановив загальний вираз для функції $Q(\tau/x_0, t_0)$ розподілу тривалості викидів у вигляді повільно збіжного ряду, членами якого є інтеграли необмежено зростаючої кратності:

$$Q(\tau|x_0, t_0) = P(\xi_{\text{в}} > \tau|x_0, t_0) = 1 - \frac{1}{f_0(x_0, t_0)} \sum_{r=1}^{\infty} (-1)^{r-1} \times \\ \times \int_{t_0}^{t_0+\tau} \int_{t_0}^{t_0+t_1} \dots \int_{t_0}^{t_0+t_{r-1}} f_r(x_0, t_0, x_0, t_1, \dots, x_0, t_r) dt_1, \dots, dt_r; \quad (1.8)$$

де

$$f_r(x_0, t_0, x_0, t_1, \dots, x_0, t_r) = \int_0^\infty \int_{-\infty}^0 \dots \int_{-\infty}^0 |y, y_1 \dots y_r| W_{2(r+1)}(x_0, y, x_0, y_1 \dots x_0, y_r, t_0, t_1, \dots, t_r) dy_1 \dots dy_r dy;$$

$$f_0(x_0, t_0) = \int_0^\infty y W_2(x_0, y, t_0) dy,$$

$W_{2(r+1)}(\dots)$ – спільна густина розподілу ймовірностей процесу і його похідної у моменти часу t_0, t_1, \dots, t_r . Формули, аналогічні (1.8), отримуються для розподілу тривалості інтервалів між викидами [11] і для густини розподілу часу першого досягнення заданої границі [11,17]. Обчислення навіть тільки перших декількох членів ряду (1.8) наштовхується на важкоподолані перешкоди, незважаючи на використання швидкодійних електронних обчислювальних машин. Якщо обмежитись наближенням першого порядку, то вираз (1.8) дає розподіл, близький до дійсного тільки за малої тривалості викидів і не дає скільки-небудь правильної, навіть якісної, уяви про цей розподіл поза інтервалом малих значень тривалості.

Порівняно нескладний вигляд мають точні формули для середнього та дисперсії числа перетинань випадковим процесом заданого рівня $x = x_0$. Середнє число перетинань з позитивною похідною на інтервалі $(t, t+T)$ дорівнює

$$\bar{n}(x_0, t, T) = \int_0^{t+T} \int_0^\infty y W_2(x_0, y, t_1) dy dt_1, \quad (1.9)$$

а дисперсія числа зазначених перетинань

$$D(x_0, t, T) = \bar{n}(x_0, t, T) - \bar{n}^2(x_0, t, T) + \int_t^{t+T} \int_t^{t+T} f_2(x_0, x_0, t_1, t_2) dt_1 dt_2, \quad (1.10)$$

де

$$f_2(x_0, x_0, t_1, t_2) = \int_0^\infty \int_0^\infty W_4(x_0, y_1, x_0, y_2, t_1, t_2) dy_1 dy_2.$$

Відносно просто обчислюється також середнє число максимумів (мінімумів) за одиницю часу стаціонарного випадкового процесу. Так, середнє число максимумів за одиницю часу, величина яких перевищує x_0 дорівнює

$$\mu_{max}(x_0) = \int_{x_0}^\infty \int_{-\infty}^0 |z| W_3(x, 0, z) dz dx,$$

де $W_3(x, y, z)$ – тривимірна густина розподілу ймовірностей процесу та його перших двох похідних у співпадаючі моменти часу.

Для ергодичних процесів відносно легко знаходяться середні тривалості викидів та інтервалів між ними

$$\mathcal{M}\{\xi_B\} = \frac{1-F_1(x_0)}{\lambda(x_0)}, \quad \mathcal{M}\{\xi_i\} = \frac{F_1(x_0)}{\lambda(x_0)},$$

де $F_1(x_0)$ – одновимірна інтегральна функція розподілу процесу, $\lambda(x_0) = (x_0)$ – середнє число перетинів за одиницю часу рівня $x = x_0$ із заданим знаком похідної

$$\lambda(x_0) = \int_0^\infty y W_2(x_0, y) dy. \quad (1.11)$$

Останніми роками, замість прагнення отримати придатні для практичних розрахунків точні формули, дослідження функцій розподілу викидів іде, як по шляху створення наближених методів визначення зазначених функцій, так і проведення широких експериментів з метою одержання емпіричних розподілів [1]. Ефективним методом наближеного визначення ймовірнісних характеристик викидів стаціонарних випадкових процесів є їх апроксимація стаціонарними випадковими послідовностями з використання ймовірнісних характеристик викидів послідовностей у якості наближення до характеристик викидів вихідного процесу. Точність апроксимації визначається вибором інтервалу дискретизації випадкового процесу. Проілюструємо ідею цього методу на найпростішому прикладі, а саме заміною випадкового процесу однозв'язною послідовністю Маркова випадкових величин ξ_n . Імовірність перевищення граничного рівня серії із k значень послідовності ξ_n, \dots, ξ_{n+k} дорівнює

$$P_1(k) = P_\sim^{k-1} (1 - P_\sim), \quad k = 1, 2, \dots, \quad (1.12)$$

де $P_\sim = P\{\xi_m > \frac{x_0}{\xi_{m-1}} > x_0\}$.

Якщо інтервал дискретизації стаціонарного випадкового процесу $\xi(t)$ дорівнює τ_0 , то

$$P_\sim = \frac{P\{\xi(t) > x_0, \xi(t+\tau_0) > x_0\}}{P\{\xi(t) > x_0\}} = \frac{\int_{x_0}^\infty \int_{x_0}^\infty W_2(x_1, x_2, \tau_0) dx_1 dx_2}{\int_{x_0}^\infty W_1(x) dx}. \quad (1.13)$$

Так, наприклад, для нормального випадкового процесу з нульовим середнім значенням і функцією кореляції $R(\tau)$ при $x_0 = 0$ з (1.13) випливає

$$P_\sim = 1 - \frac{1}{\pi} \arccos R(\tau_0).$$

У цьому випадку з (1.12) одержуємо наступну апроксимацію функції розподілу тривалості викидів над нульовим рівнем стаціонарного нормального випадкового процесу

$$P_1[(k-1)\tau_0 < \tau < k\tau_0] = \frac{1}{\pi^k} [\pi - \arccos R(\tau_0)]^{k-1} \arccos R(\tau_0).$$

Більш точна апроксимація виходить при використанні багато зв'язних моделей для апроксимуючої послідовності. Також метод успішно використовується і для наближеного обчислення розподілу часу першого досягнення границі [17].

2 ВИПАДКОВІ ПРОЦЕСИ У НЕЛІНІЙНИХ ІНЕРЦІЙНИХ СИСТЕМАХ

Аналітичне дослідження випадкових процесів у нелінійних інерційних системах (а всяка реальна система є інерційною) наштовхується на великі труднощі. Це пояснюється тим, що при аналізі нелінійних систем необхідно визначати не тільки математичні очікування та функції кореляції випадкових процесів, але й закони їх розподілу. Наявність же в системі інерційних ланок не дозволяє, у загальному випадку, знайти точний закон розподілу. Тому при аналітичному дослідженні випадкових процесів у нелінійних системах доводиться, як правило, застосовувати наближені методи, засновані або на лінеаризації системи, або на припущенні, що процеси в системі мають нормальний закон розподілу або можуть уважатися процесами Маркова невисокого порядку. Деякі з методів, наприклад, розглянутий нижче метод статистичної лінеаризації, вимагають одночасно як лінеаризації системи, так і припущення про нормальність законів розподілу. Найпоширенішими на даний час є наступні методи аналізу нелінійних інерційних систем [6,17]:

- метод лінеаризації;
- метод статистичної лінеаризації;
- метод марковських процесів, заснований на застосуванні рівнянь Фоккера–Планка–Колмогорова.

Нижче дається коротка характеристика кожного із цих методів.

2.1 Метод лінеаризації

Припустимо, що розглядається динамічна система, яка складається із сукупності m нелінійних безінерційних елементів з характеристиками $\varphi_1(x), \dots, \varphi_m(x)$ та m лінійних інерційних ланок з передатними функціями $K_1(s, t), \dots, K_m(s, t)$ і містить коло зворотного зв'язку рис. 2.1.

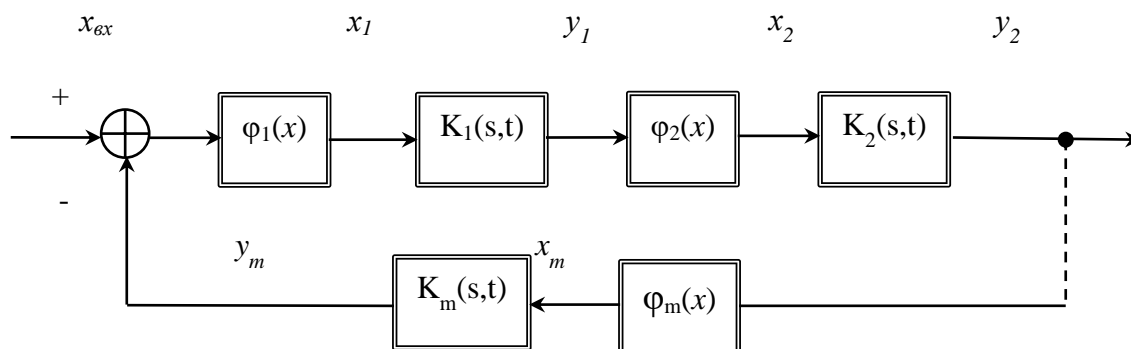


Рисунок 2.1– Структурна схема нелінійної інерційної системи з колом зворотного зв'язку

Метод лінеаризації полягає в тому, що характеристика $\varphi_i(x)$ кожного нелінійного елемента на своїй робочій ділянці апроксимується прямою лінією, що проходить через робочу точку і є дотичною у цій точці до реальної характеристики. При цьому система, у цілому, виявляється лінійною і може досліджуватися одним з методів, описаних у попередньому розділі. Очевидно, що метод лінеаризації доцільно застосовувати лише для порівняно невеликих відхилень від робочих точок, тобто при невеликій інтенсивності коливань. У деяких випадках, наприклад, при наявності на робочій ділянці зони нечутливості цей метод неприйнятний навіть при досить-таки малих значеннях інтенсивності коливань.

2.2 Метод статистичної лінеаризації

Цей метод, як і попередній, оснований на заміні кожного нелінійного елемента системи відповідним лінійним еквівалентом. Однак метод статистичної лінеаризації застосовується лише у тих випадках, коли описана вище проста лінеаризація неприйнятна, наприклад, при великій інтенсивності коливань або при наявності на робочій ділянці зони нечутливості. Тому при заміні реального нелінійного елемента лінійним виходять із міркувань рівнозначності обох елементів у статистичному сенсі, тобто з погляду однаковості статистичних характеристик процесів на їхніх виходах.

Нехай процеси на вході та виході реального нелінійного елемента з характеристикою $y = \varphi(x)$ є $x(t)$ та $y(t)$, відповідно (рис. 2.2).



Рисунок 2.2– Схема статистичної лінеаризації нелінійного елемента

При заміні цього елемента його лінійним еквівалентом намагаються забезпечити еквівалентність повних статистичних характеристик процесів $z(t)$ та $y(t)$, тобто їхніх багатовимірних розподілів. Однак, очевидно, що ніяка лінійна система не може бути повністю у всіх відносинах еквівалентна нелінійній системі. Тому доводиться обмежуватися вимогою збігу не усіх, а лише деяких статистичних характеристик процесів $z(t)$ і $y(t)$.

Найпоширенішими є наступні критерії статистичної еквівалентності лінійного еквіваленту та нелінійного елемента:

1- й критерій. Повинні збігатися математичні очікування і дисперсії процесів $z(t)$ і $y(t)$

$$\begin{aligned} \overline{z(t)} &= \overline{y(t)}, \\ \sigma_z^2(t) &= \sigma_y^2(t). \end{aligned} \quad (2.1)$$

2- й критерій. Середньоквадратичне відхилення процесу $z(t)$ від $y(t)$ повинне бути мінімальним:

$$\overline{[z(t) - y(t)]^2} = \min.$$

При використанні кожного із цих критеріїв передбачається, що лінійний еквівалент є лінійною безінерційною ланкою, з коефіцієнтами передачі для детермінованої та флуктуаційної складових вхідного процесу $x(t)$, рівними відповідно k_0 и k_1 . Очевидно, що детермінована складова вхідного процесу співпадає з його математичним сподіванням $m = \overline{x(t)}$, а флуктуаційна складова являє собою відцентрований випадковий процес

$$x^0(t) = x(t) - m(t).$$

$$\text{Тому} \quad z(t) = k_0 m(t) + k_1 x^0(t). \quad (2.2)$$

Із співвідношень (2.1) – (2.2) неважко встановити, що при застосуванні першого критерію еквівалентності, параметри лінійного еквівалента визначаються наступними виразами:

$$k_0 = \sqrt{\frac{\bar{y}}{\bar{x}}}; \quad k_1 = \pm \sqrt{\frac{\sigma_x^2}{\sigma_y^2}}, \quad (2.3)$$

при цьому знак плюс перед радикалом береться у тих випадках, коли характеристика $y=\varphi(x)$ у межах робочої ділянки зростає і знак мінус, якщо вона убуває.

При застосуванні другого критерію замість (2.3) виходить

$$k_0 = \sqrt{\frac{\bar{y}}{\bar{x}}}; \quad k_1 = \frac{\overline{x^0 y}}{\sigma_x^2}$$

Надалі, заради стислості викладок, обмежимося розглядом співвідношень (2.3). Із цих співвідношень випливає, що для визначення параметрів k_0 і k_1 лінійного еквіваленту необхідно знати математичні очікування та дисперсії процесів на вході та виході нелінійного елемента. Оскільки, $y=\varphi(x)$, то

$$\bar{y} = \int_{-\infty}^{\infty} W(x)\varphi(x)dx$$

$$\sigma_y^2 = \int_{-\infty}^{\infty} W(x)\varphi^2(x)dx - \bar{y}^2, \quad (2.4)$$

де $W(x)$ – одномірний закон розподілу процесу $x(t)$ на вході нелінійного елемента. Припустимо, що параметри лінійних еквівалентів усіх нелінійних елементів системи (рис. 2.1) відомі. Тоді розрахунки математичних очікувань та кореляційних функцій, а отже і дисперсій процесів у будь-якій точці цієї системи не представляють труднощів принципового характеру, якщо система може вважатися стаціонарною. Дійсно, при заміні усіх

нелінійних елементів їх лінійними еквівалентами система, у цілому, та усі її елементи виявляються лінійними. Тому, якщо представити процеси в усіх точках системи у вигляді суми їх математичних сподівань $\overline{x_i(t)}$ та флуктуаційних (центрованих) складових $x_i^0(t)$, то можна виконувати аналіз системи незалежно, спочатку для математичних сподівань процесів, а потім для їх флуктуаційних складових, використовуючи при цьому значення коефіцієнтів передачі лінійних еквівалентів для математичних очікувань і для флуктуаційних складових, відповідно. При цьому годиться будь-який метод аналізу, прийнятний для аналізу лінійних систем. Як уже відзначалося, знаходження математичних сподівань та функцій кореляції у будь-яких точках лінійної стаціонарної системи не являє собою принципових труднощів. Однак, у випадку стаціонарної лінеаризованої системи є труднощі, причому принципового характеру, суть яких полягає у наступному. За припущеннями параметри k_0 і k_1 кожного лінійного еквіваленту відомі, тоді як у дійсності вони невідомі і для їхнього обчислення потрібне знання законів розподілу $W_i(x)$ процесів на входах усіх нелінійних елементів. Але визначити закони розподілу процесів у різних точках інерційної системи, у загальному випадку, не вдається, навіть якщо ця система лінійна. Для подолання цих труднощів при застосуванні методу статистичної лінеаризації роблять ще одне припущення – процеси на входах усіх нелінійних елементів мають нормальний закон розподілу. При цьому необхідний для обчислення по формулах (2.4) одномірний розподіл $W(x)$ повністю визначається значеннями \bar{x} і σ_x^2 , а саме

$$W(x) = \frac{1}{\sqrt{2\pi} \sigma_x} \exp \left[-\frac{(x-\bar{x})^2}{2\sigma_x^2} \right], \quad (2.5)$$

і обчислення параметрів k_0 , k_1 уже неважко виконати. Проте, труднощі практичного характеру можуть бути істотними, якщо система має щонайменше один нелінійний елемент, але охоплений зворотним зв'язком. Дійсно, унаслідок наявності зворотного зв'язку значення математичного сподівання та дисперсії процесу на вході нелінійного елемента залежать від параметрів k_0 і k_1 його лінійного еквіваленту:

$$\bar{x} = f_1(k_0, k_1), \quad \sigma_x^2 = f_2(k_0, k_1).$$

Параметри k_0 і k_1 , як це випливає зі співвідношень (2.3) – (2.5) у свою чергу залежать від \bar{x} і σ_x^2

$$\begin{aligned} k_0 &= F_1(\bar{x}, \sigma_x^2), \\ k_1 &= F_2(\bar{x}, \sigma_x^2). \end{aligned} \quad (2.6)$$

Тому для знаходження параметрів k_0 і k_1 рівняння (2.6) необхідно розв'язувати спільно. Точний аналітичний розв'язок цих рівнянь вдається отримати лише у найпростіших, окремих випадках. У більш складних випадках, вищезазначені рівняння виявляються трансцендентними, і їх розв'язок

знаходять графічно. У ще більш складних випадках доводиться застосовувати методи послідовних наближень [11].

Незважаючи на те, що метод статистичної лінеаризації заснований на цілій низці порівняно грубих припущень і вимагає, у багатьох випадках, трудомістких обчислень, він є на сьогодні єдиним методом, який дозволяє визначати характеристики випадкових процесів у суттєво нелінійній (де проста лінеаризація неприпустима) інерційній системі високого порядку ($n > 2$) без застосування обчислювальних машин.

Однак, у деяких випадках припущення, на яких заснований метод статистичної лінеаризації, виявляються занадто грубими. Зокрема, цей метод не гарантує близькості спектра процесу на виході лінійного еквіваленту до спектру на виході нелінійного елемента, наприклад, у [16] зазначається, що у випадку коли спектр вхідного сигналу зосереджений в основному поблизу частоти ω_0 , а на виході нелінійного елемента увімкнено фільтр, що має резонанс на частоті $2\omega_0$ або $3\omega_0$ і т.д., то можуть виникати неприпустимі помилки. Тому зазвичай, результати отримані методом статистичної лінеаризації, намагаються надалі уточнювати більш строгими аналітичними методами (при $n \leq 2$), або шляхом застосування математичного моделювання (при $n > 2$).

2.3 Застосування рівняння Фоккера – Планка – Колмогорова

При дослідженні процесів на виході нелінійних систем, які описують нелінійними диференціальними рівняннями, у деяких випадках може бути ефективно використаний математичний апарат теорії марковських процесів та дифузійних рівнянь. Проілюструємо це на прикладі нелінійного диференціального рівняння першого порядку

$$\frac{d\eta(t)}{dt} = f[\eta(t), \xi(t)], \quad (2.7)$$

де $\eta(t)$ – випадковий процес на виході системи, а $\xi(t)$ – випадковий процес на її вході. Математичний апарат марковських процесів можна застосовувати до рівняння (2.7) у тому випадку, коли час кореляції τ_k вхідного процесу значно менший часу зміни T_{ex} вихідного процесу ($\tau_k \ll T_{ex}$), який приблизно визначається виразом:

$$T_{вх} \sim \left[\frac{\partial f}{\partial \eta} \right]^{-1}$$

За такої умови вихідний процес $\eta(t)$ є приблизно марковським і описується дифузійними рівняннями при відповідному виборі коефіцієнтів $A(x, t)$, $B(x, t)$. Ці коефіцієнти знаходять за допомогою формул [6]:

$$A(x, t) = \overline{f[x, \xi(t)]} + \int_{-\infty}^0 \left\{ \frac{\partial f[x, \xi(t)]}{\partial \eta} f[x, \xi(t + \tau)] - \frac{\partial f[x, \xi(t)]}{\partial \eta} \overline{f[x, \xi(t + \tau)]} \right\} d\tau;$$

$$B(x, t) = 2 \int_0^\infty \{ \overline{f[x, \xi(t)]f[x, \xi(t + \tau)]} - \overline{f[x, \xi(t)]} \overline{f[x, \xi(t + \tau)]} \} d\tau, \quad (2.8)$$

де значення $x = \eta(t)$ у правій частині (2.8) вважаються не випадковими.

До рівняння типу (2.7) приводиться, наприклад, нелінійне стохастичне рівняння, яке описує найпростішу модель системи автоматичного підстроювання частоти. Позначаючи через $\varphi(t)$ розходження фаз керованого генератора і гармонійного сигналу керування, можна отримати рівняння

$$\dot{\varphi}(t) = \omega - \omega_0 - k[A_0 \sin \varphi(t) + \xi(t)], \quad (2.9)$$

де ω_0 - власна частота генератора; A_0, ω - амплітуда та частота сигналу керування; k - деяка константа; $\xi(t)$ - нормальний білий шум зі спектральною густиною N_0 . Обчислення за формулами (8.8) у випадку рівняння (2.9) проводяться відносно просто і приводять до наступних коефіцієнтів дифузійного рівняння:

$$A(\varphi) = \omega - \omega_0 - kA_0 \sin \varphi, \quad B(\varphi) = k^2 N_0. \quad (2.10)$$

Застосування рівняння Фоккера-Планка-Колмогорова дозволяє вирішувати наступні задачі:

- знаходити одномірний стаціонарний розподіл ймовірностей на вході нелінійних систем першого порядку;

- для процесу синхронізації генераторів за наявності завад, який пов'язаний з рівнянням (2.9), знаходити як стаціонарний розподіл по φ , так і відхилення від ідеального синхронного режиму, тобто підраховувати обумовлене завадами середнє число зривів синхронізації за точними формулами [11,17];

- розглядати питання, пов'язані з досягненням меж, досліджувати стійкість різних систем за наявності завад, ймовірності зривів стеження за заданий час та тощо, і при цьому можна користуватися як точними результатами, отриманими за допомогою оберненого рівняння Фоккера-Планка-Колмогорова, а також наближеними формулами [11];

- застосовуючи більш складні методи (малого параметра або розв'язання у ряди), вирішувати низку інших задач, наприклад обчислення кореляційної функції на виході нелінійного пристрою, задачі, пов'язані з багатомірними рівняннями та тощо [11].

Проілюструємо сказане на прикладі розв'язку задачі про синхронізацію генераторів за наявності завад. Коефіцієнтам (2.10) відповідає рівняння Фоккера-Планка-Колмогорова

$$\dot{W}(\varphi) = - \left[\frac{\partial [(\omega - \omega_0 - kA_0 \sin \varphi) W(\varphi)]}{\partial \varphi} \right] + (1/2) k^2 N_0 \frac{\partial^2 W(\varphi)}{\partial \varphi^2}. \quad (2.11)$$

Стаціонарний розподіл $W_{ст}(\varphi)$ задовольняє умові $\dot{W}(\varphi) = 0$, у силу справедливості (2.10) рівнянню

$$-\frac{\partial[(\omega-\omega_0-kA_0\sin\varphi)]W_{\text{ст}}(\varphi)}{\partial\varphi} + (1/2)k^2N_0\frac{\partial^2W_{\text{ст}}(\varphi)}{\partial\varphi^2} = 0. \quad (2.12)$$

Інтегруючи останнє рівняння перший раз, отримуємо

$$(\omega - \omega_0 - kA_0\sin\varphi)W_{\text{ст}}(\varphi) - (1/2)k^2N_0\frac{\partial W_{\text{ст}}(\varphi)}{\partial\varphi} = C_1,$$

де константа C_1 інтерпретується як потік ймовірності через перетин $\varphi(t) = \text{const}$. Коли б режим синхронізації був ідеальним, то потік ймовірності C_1 був би нульовим. Ненульове значення C_1 означає наявність ненульового значення середньої похідної

$$\bar{\dot{\varphi}} = 2\pi C_1. \quad (2.13)$$

Якщо виконати повторне інтегрування рівняння (2.11) та ввести другу постійну інтегрування C_2 , то знаходимо

$$W_{\text{ст}}(\varphi) = e^{D\varphi + D_c\cos\varphi} \left[C_2 - \frac{2C_1}{k^2N_0} \int_0^\varphi e^{-D\theta - D_c\cos\theta} d\theta \right], \quad (2.14)$$

де введено наступні позначення

$$D = \frac{2(\omega - \omega_0)}{k^2N_0}, \quad D_c = \frac{2A_0}{kN_0}.$$

Постійні інтегрування C_1, C_2 відшукуються шляхом задоволення умовам періодичності та нормування до одиниці

$$W_{\text{ст}}(2\pi) = W_{\text{ст}}(0), \quad \int_0^{2\pi} W_{\text{ст}}(\varphi) d\varphi = 1.$$

Тим самим повністю розраховується стаціонарна густина розподілу ймовірностей (2.14) та обумовлений завадами зсув частоти (2.13). Після проведення обчислень, як показано в [11], останньому результату можна надати вигляд

$$\bar{\dot{\varphi}} = (\omega - \omega_0) \frac{\text{sh } \pi D}{\pi D} |I_{iD}(D_c)|^{-2},$$

де $I_{iD}(D_c)$ — функція Бесселя уявного аргументу і уявного індексу.

Наостанок, наведемо у якості прикладу, відповідно до п. 3, розрахунки зривів процесу стеження для деякої системи автоматичного регулювання. Її структурна схема зображена на рис. 2.3. Нелінійний елемент має ділянки характеристики, які спадають; скажемо $F(\eta)$ убуває при $|\eta| > \eta_0$ і зростає при $|\eta| < \eta_0$ і має властивість $F(-\eta) = -F(\eta)$. Система здійснює спостереження за флуктуаційним сигналом $\zeta(t)$.

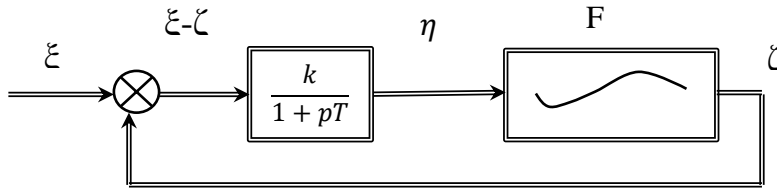


Рисунок 2.3 - Замкнена нелінійна система стеження

Зазначеній схемі, відповідає рівняння

$$\eta = \frac{k}{1 + pT} (\xi - \zeta); \quad \zeta = F(\eta).$$

Якщо у цьому рівнянні позбавитись ζ , то отримаємо рівняння

$$\dot{\eta} = -\frac{1}{T}\eta + \frac{k}{T} [\xi(t) - F(\eta)],$$

яке є окремим випадком (2.7). За формулою (2.8) знаходимо коефіцієнти зносу та дифузії

$$A(x, t) = -\frac{x}{T} - \frac{k}{T} F(x;),$$

$$B(x, t) = \frac{k^2}{T^2} \int_{-\infty}^{\infty} \overline{\xi(t)\xi(t+\tau)} d\tau, \quad \text{при } \overline{\xi(t)} = 0. \quad (2.15)$$

Нехай $\pm x_1$ — точки, у яких коефіцієнт зносу A змінює знак:

$$A < 0 \quad \text{при } x - x_1, 0 < x < x_1; \quad A > 0 \quad \text{при } -x_1 < x < 0.$$

Коли $\eta(t)$ досягне значень $\pm x_1$, то у половині випадків точка $\eta(t)$ повернеться у область $|x| < x_1$, а у половині випадків потрапить у область неповернення, тобто $|x| > x_1$, що відповідає зриву процесу стеження. Щоб знайти середню частоту λ зривів, потрібно обчислити середній час M досягнення меж $\pm x_1$, тоді

$$\lambda = \frac{1}{2M}. \quad (2.16)$$

Середній час M можна знайти, розглядаючи наступне рівняння за нульових умов на межі:

$$W(\pm x_1, t, x_0, t_0) = 0.$$

Інший шлях полягає у знаходженні середнього часу досягнення меж із точки x_0 , тобто функції

$$M(x_0) = - \int_{-t_0}^{\infty} (t - t_0) d \int_{-x_1}^{x_1} W(x, t, x_0, t_0) dx \quad (2.17)$$

Оскільки коефіцієнти

$$A(x, t) = A(x), \quad B(x, t) = B(x)$$

не залежать від t , то з нього можна вивести рівняння

$$\frac{1}{2}B(x_0) \frac{d^2 M(x_0)}{dx_0^2} + A(x_0) \frac{dM(x_0)}{dx_0} + 1 = 0, \quad (2.18)$$

яке слід розв'язувати за нульових граничних умов $M(\pm x_1)=0$. Рівняння (2.18) є лінійним відносно $\frac{dM(x_0)}{dx_0}$ і може бути вирішене точно. Враховуючи симетрію функції $M(x_0)$, знаходимо його розв'язок із зазначеними умовами

$$M(x_0) = -\frac{2}{B} \int_{-x_1}^{x_1} dy \int_0^y e^{U(y)-U(z)} dz, \quad (2.19)$$

$$\left(\frac{dU(x)}{dx} = -\frac{2A(x)}{B} \right).$$

У припущенні, що зриви стеження відбуваються відносно рідко, результати виглядають особливо простими. При цьому формулу (2.16) можна записати:

$$\lambda = \frac{1}{2M(0)},$$

а моменти часу, за яких трапляються зриви стеження утворюють потік приблизно пуассонівський, так що ймовірність відсутності зриву за час t визначається формулою

$$P(t) = e^{\lambda t}.$$

Формула (2.19) при цьому значно спрощується. За умови, що величина коефіцієнту дифузії B відносно мала, будемо мати наближений результат

$$M(0) \approx \frac{2}{B} \int_{-x_1}^0 e^{U(y)} dy \int_{-x_1}^0 e^{-U(z)} dz. \quad (2.20)$$

Скориставшись розвиненням у ряд отримуємо

$$U \approx U + \frac{1}{2} \ddot{U}(-x_1)(y+x_1)^2 = U(x_1) - \frac{\dot{A}(-x_1)}{B} (y+x_1)^2,$$

$$U(z) \approx U(0) + \frac{1}{2} \ddot{U}(0)z^2 = U(0) - \frac{\dot{A}(0)}{B} z^2.$$

У точках, де підінтегральні вирази максимальні, врахування виразів (2.16), (2.20) дає наближений результат для інтенсивності потоку Пуассона

$\lambda = \frac{1}{\pi} |\dot{A}(0)\dot{A}(x_1)|^{1/2} e^{U(0)-U(x_1)}$. Методи отримання наближених результатів інших типів докладніше викладені у [17].

2.4 Ймовірнісне моделювання радіотехнічних систем на електронних обчислювальних машинах

Низка проблем, пов'язаних з дослідженням та конструюванням радіотехнічних систем, розглянутих у попередніх розділах, часто не піддається точному аналітичному розв'язку і вимагає застосування чисельних методів. Так, підрахунок ймовірності виявлення сигналу, обчислення ефективності оцінок параметрів сигналу пов'язані з кратним інтегруванням по складним за конфігурацією областями, які виконуються за кубатурними формулами [5], у тому числі і з випадковими вузлами [5, 6]. Дослідження нелінійних замкнених систем вимагає чисельного інтегрування диференціальних рівнянь [4]. При синтезі оптимальних систем доводиться використовувати ітераційні схеми мінімізації складних функціоналів. Якщо оперувати з чисто математичними об'єктами (інтеграл, система рівнянь), «класичні» чисельні методи часто приховують фізичний зміст розглянутих процесів і не мають явної системної інтерпретації проміжних етапів розрахунків, не дозволяють осмислити причини підсумкової поведінки системи. Більш глибокими, у цьому сенсі, виявляється методи прямого ймовірнісного моделювання [5, 6], яке часто називають методом Монте-Карло. Він зводиться до генерації послідовності реалізацій вихідних випадкових процесів, до функціонального (алгоритмічного) їх перетворення у відповідності до заданої структури системи, що підлягає моделюванню і, нарешті, до статистичної обробки послідовності реалізацій на виході системи.

Тут аналітичне дослідження математичної моделі замінюється експериментом з фізичною моделлю, відтвореної на цифровій або аналоговій обчислювальній машині. Навіть найпростіша схема моделювання, у більшості випадків, повністю відповідає схемі натурального експерименту з радіотехнічним пристроєм, наприклад: генератор стандартних сигналів замінюється генератором випадкових діянь, досліджувана на стенді схема — відповідною програмою, вимірювальний прилад — програмою статистичної інтерпретації результатів. Враховуючи, що дорожнеча та унікальність сучасних радіотехнічних систем, найчастіше, виключають можливість повного натурального експерименту з макетом системи, залишається вважати, що моделювання є не тільки найбільш природним і наочним, але і єдиним універсальним засобом інженерних розрахунків радіотехнічних систем та комплексів [5].

Моделювання на аналогових ЕОМ зручно, якщо досліджувана система описується звичайними диференціальними рівняннями і не містить розгалуженої логіки, а вимоги до точності розв'язку відносно невисокі, порядку 1—3%. Паралельні процеси, тут природно вписуються у структуру звичайної аналогової ЕОМ. Багаторазове повторення моделювання, необхідне для дослідження нестационарних систем, реалізується шляхом багатократного повторення розв'язків за різних початкових умов [6]. Дещо специфічні труднощі виникають у тих випадках, коли ідеалізація досліджуваної сис-

теми вносить у її опис процеси, які фізично не можливо реалізувати, наприклад, білий шум. Як відомо, будь-який генератор шуму, що використовується у аналоговій ЕОМ, може дати шум із широким, але у принципі, обмеженим спектром. При цьому виявляється [11], що для уникнення грубих помилок слід розуміти інтеграли від випадкових функцій, що входять у опис системи, так звані стохастичні інтеграли, у сенсі Р. Л. Стратоновича [17], але не в розумінні К. Іто, як це прийнято у багатьох теоретичних дослідженнях. Збільшення точності розв'язку завдань, зручних для моделювання на аналогових ЕОМ, досягається шляхом використання цифрових аналогів і гібридних машин [5], а також трансляції програм з аналогових на цифрові ЕОМ за допомогою спеціальних програм [6]. Суттєво більші труднощі виникають при моделюванні на однопрограмною цифровій ЕОМ процесів, що паралельно протікають у часі, наприклад, у багатофазних системах масового обслуговування. У цьому випадку необхідно здійснювати періодичний огляд усіх елементів моделі або впорядковувати ці елементи за часом чергової зміни кожного їх стану [5]. Стандартизація цих операцій, у рамках різних систем автоматизації програмування, розглянута, наприклад, у [6].

Найбільш важкою (хоч і неспецифічною для ймовірнісних завдань) є проблема побудови моделі, тобто проблема «розумного спрощення» опису досліджуваної системи. Рекомендації класичної теорії подібності [5] можуть принести користь при конструюванні порівняно простих (особливо аналогових) моделей. У складних випадках основну роль відіграє інтуїція дослідника, який керується поки лише евристичними, загальними принципами і класифікацією моделей [14]. Наведемо приклади цих загальних принципів:

- компроміс між повнотою моделі, трудомісткістю побудови моделі та складністю експериментування з нею;
- відповідність точності окремих елементів моделі та точності вихідних даних для моделювання;
- розбивка процесу моделювання на ряд незалежних етапів;
- точне відтворення одних частин системи, та спрощене – для інших;
- виключення окремих частин системи із заміною їх еквівалентними випадковими діями;
- створення послідовності моделей наростаючої складності з перевіркою збіжності результатів – метод Рунге.

Мабуть, однією з найближчих перспектив стандартизації та автоматизації процесу побудови моделі, що розглядається, як окремий випадок процесу абстрагування, може бути евристичне програмування [5, 6].

3 МАТЕМАТИЧНИЙ СИНТЕЗ ОПТИМАЛЬНИХ РАДІОТЕХНІЧНИХ СИСТЕМ

3.1 Базові положення теорії прийняття статистичних рішень

Застосування математичних методів, що відносяться до загальної теорії оптимальних систем та математичного програмування, дозволяє знайти аналітично такі перетворення (алгоритми), які потрібно зробити над вхідними процесами та сигналами, щоб на виході системи отримати результати, найкращі у певному, строго зафіксованому сенсі. Після цього радіотехнічними пристроями, засобами автоматики або обчислювальної техніки залишається здійснити ці перетворення практично, найчастіше у формі перетворень електричних сигналів. Така система буде оптимальною. Звичайно, практичне здійснення зазначених математичних перетворень може виявитися не простою задачею, однак пов'язані із цим технічні труднощі у даному навчальному посібнику не розглядаються. Не розглядаються також питання оцінки складності оптимальних алгоритмів-перетворень і вартості відповідних систем. При застосуванні математичного апарату для розв'язку інженерних задач, які можуть бути досить різноманітними, кожна задача повинна бути чітко формалізована. Із низки різноманітних інженерних вимог до системи, повинен бути створений один – основний критерій оптимальності, а саме – фіксована функція втрати. Обраний критерій може бути комбінованим, тобто загальна функція втрат може складатися з елементарних функцій втрат або бути їх комбінацією. Відповідно до цього, складові частини системи, оптимальної відносно комбінованого критерію, часто можна отримувати шляхом розгляду більш простих задач – оптимізації підсистеми за окремим елементарним критерієм. Окрім виду оптимальних перетворень, які визначають структуру оптимальної системи, становить інтерес відповідна до цих перетворень величина середніх втрат, що характеризує якісні показники синтезованої системи. У зв'язку з цим бажано мати математичні методи, що дозволяють обчислювати цю величину.

Спочатку стисло викладемо основні, базові положення простої однокрокової задачі прийняття статистичних рішень, яку можна розглядати як окрему частину загальної теорії оптимізації динамічних систем. Припустимо, що необхідно оцінити (виміряти, відтворити) невідому величину або декілька величин θ за умови, що спостерігається інша випадкова величина або декілька величин y , функціонально пов'язаних з θ . Процес спостереження y може являти собою спотворений завадами корисний (інформативний) сигнал $S(\theta)$, у якому «закодований» інформаційний параметр (або параметри) θ . Так, у низці радіотехнічних задач спостереженню підлягає адитивна суміш сигналу $S(\theta)$ і шуму $n(t)$. Тут і надалі під $y(t)$ та $S(\theta, t)$ слід розуміти сукупність векторів заданої розмірності або реалізацій процесів $y(t)$ та $S(\theta, t)$ на певному часовому інтервалі.

У більш загальному випадку, процес спостережень описують математичною моделлю

$$y(t) = S(\theta, \varphi, t) + n(t),$$

де $n(t)$ – адитивна завада; θ – корисний, інформативний параметр; φ — супутній, неінформативний випадковий параметр, що заважає процесу спостереження. У низці випадків $\theta, \varphi \in$ функціями часу. Статистичний опис процесу спостережень $y(t)$ будемо вважати заданим у вигляді умовного розподілу ймовірностей $\mathcal{W}(y/\theta)$. Якщо розподіл $\mathcal{W}(y/\theta)$ розглядають як функцію від параметра θ , то його прийнято називати функцією правдоподібності [3]. Потрібно знайти оптимальне перетворення (правило, алгоритм), який дозволяє на основі результатів спостережень за величиною y знаходити оптимальну оцінку $\theta^*(y)$ невідомого інформативного параметра θ . Критерій оптимальності пов'язують із, так званою, функцією вартості $C(\theta^*, \theta)$ (інакше – функцією збитків, втрат або штрафів). Тип її вибирається на підставі інженерних або інтуїтивних міркувань. Так, коли значення θ^* повинне бути якнайближчим до дійсного значення θ , то доцільно вибрати $C(\theta^*, \theta)$ з мінімумом θ^* поблизу точки θ – істинного значення інформативного параметру, яке до речі невідоме. Конкретне значення $C(\theta^*, \theta)$ при $\theta^* = \theta^*(y)$ виявляється випадковим, тому якість системи потрібно визначати деякими середніми показниками. Осередненням $C(\theta^*, \theta)$ при $\theta^* = \theta^*(y)$ отримуємо, так званий умовний ризик

$$\mathcal{R}(\theta^*/\theta) = \int_{(y)} C[\theta^*(y), \theta] \mathcal{W}(y/\theta) dy. \quad (3.1)$$

Зазвичай, більш кращими правилами (алгоритмами) оцінювання інформативного параметра $\theta^*(y)$ вважаються ті, які призводять до менших значень втрат $\mathcal{R}(\theta^*/\theta)$. Якщо деяке $\theta^*(y)$ мінімізує $\mathcal{R}(\theta^*/\theta)$ для усіх значеннях θ , то таке правило (розв'язок) є оптимальним. Однак, у загальному випадку, розв'язок, який доставляє мінімум $\mathcal{R}(\theta^*/\theta)$, буде різним при різних θ . У цьому випадку вибір оптимального правила можна робити на основі різних підходів: байєсового, мінімаксного, Неймана–Пірсона та ін. Наприклад, у байєсовій теорії вважається заданим апіорний розподіл $\mathcal{W}(\theta)$ невідомого інформативного параметра θ , і тоді розглядається середній безумовний ризик, пов'язаний з вибраним правилом оцінювання

$$\mathcal{R}(\theta^*) = \int_{(\theta)} C(\theta^*, \theta) \mathcal{W}(\theta) d\theta = \int_{(\theta)} \int_{(y)} C[\theta^*(y), \theta] \mathcal{W}(\theta, y) d\theta dy. \quad (3.2)$$

Враховуючи, що $\mathcal{W}(\theta, y) = \mathcal{W}(y) \mathcal{W}(\theta/y)$, де $\mathcal{W}(\theta/y)$ — апостеріорний розподіл параметра θ , формула (3.2) може бути записана у вигляді

$$\mathcal{R}(\theta^*) = \int_{(y)} \mathcal{R}[\theta^*(y)/y] \mathcal{W}(y) dy, \quad (3.3)$$

де $\mathcal{R}[\theta^*(y)/y] = \int_{(\theta)} C(\theta^*, \theta) \mathcal{W}(\theta/y) d\theta$ – умовний по y середній ризик.

Правило $\theta^*(y)$, яке мінімізує середній ризик (3.3), називають оптимальним байєсовим правилом або розв'язком. З формули (3.3) видно, що мінімізація по y середнього ризику $\mathcal{R}(\theta^*)$ зводиться до мінімізації по θ виразу для умовного середнього унаслідок позитивності $\mathcal{W}(\theta)$:

$$\mathcal{R}[\theta^*(y)/y] = \int_{(\theta)} C(\theta^*, \theta) \mathcal{W}(\theta/y) d\theta \quad (3.4)$$

$$\int_{(\theta)} C(\theta^*, \theta) \mathcal{W}(\theta/y) d\theta = \min_{\theta^*} \quad (3.5)$$

Задачу (3.5) часто можна розв'язати аналітично.

У деяких випадках апіорний розподіл $\mathcal{W}(\theta)$ невідомого інформаційного параметра не можна вважати заданим. Тоді застосовують мінімакський критерій, тобто мінімізують умовний ризик для найбільш несприятливого випадку, а саме – знаходять оптимальне мінімаксне правило оцінювання за умови

$$\min_{\theta^*} \max_{\theta} \mathcal{R}(\theta^*, \theta) = \max_{\theta} \mathcal{R}(\theta_{min}^*, \theta). \quad (3.6)$$

Мінімакський критерій (3.6) дає оптимальний розв'язок лише для найгіршого випадку. Він рідше використовується у прикладних дослідженнях, ніж байєсовий критерій, тому що аналітичні дослідження тут пов'язані з більшими труднощами. Критерій Неймана – Пірсона буде розглянутий у наступному підрозділі, окрім того низка подальших підрозділів містить приклади застосування, викладених вище, базових положень теорії прийняття статистичних рішень до різноманітних задач синтезу радіоелектронних пристроїв та систем.

3.2 Задачі виявлення та розпізнавання сигналів

Задача виявлення сигналу являє собою задачу прийняття рішення про наявність або відсутність сигналу за результатами спостережень вхідного процесу $y(t)$. Якщо сигнал відсутній, то вхідний процес являє собою чисту заваду $n(t)$ із законом розподілу $\mathcal{W}(y/n)$, який вважається апіорі відомим. Якщо у вхідному процесі присутній корисний сигнал з параметрами θ , то цій ситуації відповідає закон розподілу $\mathcal{W}(y/\theta)$. Введемо додатковий параметр γ рівний одиниці, якщо корисний сигнал присутній у вхідному процесі, і рівний нулю, якщо – відсутній. Тоді, увівши розширений параметр $\eta = \{ \gamma, \theta \}$, можна звести задачу виявлення корисного сигналу до задачі прийняття статистичних рішень, описаної у попередньому підрозділі. Відповідно до прийнятих припущень значення $\eta = 1$ відповідає ситуації, що корисний сигнал присутній в $y(t)$, а $\eta = 0$ — сигнал відсутній. Тоді сукупність значень $\eta = 1, \gamma = 1$ відповідають правильному рішенню, так само як і $\eta = 0, \gamma = 0$. Введемо штрафи: A – за фіктивну тривогу (коли в дійсності сигнал відсутній) і B – за пропуск корисного сигналу (коли в дійсності він є), і відсутність штрафів за правильні розв'язки. Для цього випадку, функція штрафів може бути записана у такій аналітичній спосіб:

$$C(\theta, \gamma) = \begin{cases} A & \text{при } \gamma = 0, \eta = 1, \\ B & \text{при } \gamma = 1, \eta = 0, \\ 0 & \text{при } \gamma = \eta = 1, \text{ або } \gamma = \eta = 0. \end{cases} \quad (3.7)$$

Якщо апіорі відомі: густина розподілу ймовірностей параметра $\mathcal{W}(\theta)$ та імовірності $\mathcal{W}(\gamma)$ появи або відсутності сигналу, то можна застосовувати байєсове оптимальне правило прийняття статистичних рішень, за методикою викладеною у підрозділі 3.1.

Однак, часто буває так, що апіорні ймовірності $\mathcal{W}(\gamma)$, а також штрафи A, B вважаються не заданими, або доцільно так вважати. Тоді можна відшукувати правило прийняття статистичних рішень $\gamma(\mathbf{y})$, оптимальне за критерієм Неймана–Пірсона. У цьому разі, розглядаються ймовірності помилок двох типів: ймовірність P хибної тривоги та ймовірність $1-D$ – пропуску сигналу. Оптимальне правило прийняття рішень мінімізує одну ймовірність помилки при зафіксованому (заданому) значенні іншої. Як показує теорія [3], і байєсовий, і мінімаксий критерії оптимальності приводять до однієї і тієї ж процедури у процесі обробки результатів спостережень вхідного процесу. Спочатку утворюється, так зване, відношення правдоподібності

$$\Lambda(\mathbf{y}/\theta) = \frac{\mathcal{W}(\mathbf{y}/\theta)}{\mathcal{W}(\mathbf{y}/n)}. \quad (3.8)$$

При застосуванні байєсового критерію умовне відношення правдоподібності додатково осереднюють за параметром θ , тобто знаходять

$$\Lambda(\mathbf{y}) = \int_{(\theta)} \Lambda(\mathbf{y}/\theta) \mathcal{W}(\theta) d\theta. \quad (3.9)$$

Потім $\Lambda(\mathbf{y})$ або будь-яка монотонна функція від $\Lambda(\mathbf{y})$ порівнюється з деяким пороговим значенням K . При байєсовому підході величина порогу K однозначно виражається через апіорні ймовірності $\mathcal{W}(\gamma)$ і відношення штрафів A/B . Якщо ж використовується критерій Неймана–Пірсона, то поріг виражається через ймовірність хибної тривоги P або ймовірність правильного виявлення D . Залежність цих ймовірностей від K неважко отримати шляхом інтегрування:

$$\begin{aligned} D(K) &= \int_K^{\infty} P_{0,1}(\Lambda) d\Lambda, \\ F(K) &= \int_K^{\infty} P_{1,0}(\Lambda) d\Lambda, \end{aligned} \quad (3.10)$$

де $P_{0,1}(\Lambda)$ — розподіл випадкової величини Λ при $\gamma=0$ і $\gamma=1$, відповідно. Функції $D(K)$, $F(K)$ слід тлумачити, як параметричне зображення залежності D від F , і яку часто називають робочою характеристикою оптимального пристрою виявлення [2,3].

Перейдемо до задачі розпізнавання сигналів. У найпростішому випадку, ця задача зводиться до вибору однієї із двох або декількох альтернатив.

Якщо можливі лише дві альтернативи і статистичні властивості сигналів обох можливих типів відомі, то задача розпізнавання оптимально вирішується тими ж самими методами, що й задача виявлення. З принципової точки зору, так само вирішується і задача багато альтернативного байєсового розпізнавання [3]. Значно складнішою стає ситуація у тих випадках, коли апріорні розподіли заздалегідь невідомі. У цих випадках доцільно проводити попереднє «навчання», наприклад, методом показів: одночасно вказується і значення вхідної величини $y(t)$, і тип сигналу γ (для вказівки типу сигналу можна зберегти позначення $\gamma = 0; 1$). Процес «навчання» містить n показів, що дають сукупність значень $\{y_1, \gamma_1, \dots, y_n, \gamma_n\}$. Власне, етап розпізнавання полягає у подачі на систему вхідної послідовності $y(t)$, тип сигналу у якій невідомий, а від системи вимагається визначення цього типу. Шуканим є правило, що оцінює параметр γ , тобто

$$\gamma^* = \gamma [y_1, \gamma_1, \dots, y_n, \gamma_n] \quad (3.11)$$

Для розв'язання цієї задачі можна використовувати різні методи. Якщо припустити, що тип розподілу $P_\gamma(y)$ при обох значеннях γ відомий з точністю до параметрів $(\theta_1, \dots, \theta_n)$, що мають рівномірний апріорний розподіл, тоді він може бути представлений у вигляді:

$$P_\gamma(y) = \text{const} \exp\{\sum_{i=1}^k \alpha_i(\gamma) \varphi_i(y) + \varphi_0(y)\}, \quad (3.12)$$

де функції $\varphi_0(y), \dots, \varphi_k(y)$ задані (такий тип завдання зручний для розв'язування задачі, причому конкретний вибір функцій $\varphi_i(y)$ не є принциповим. У такому разі можна шукати оптимальне байєсове правило розпізнавання. При цьому у якості вхідної величини $y(t)$ слід брати усю сукупність аргументів співвідношення (3.11), а у якості величин, що підлягають визначенню — сукупність $\alpha_1^y, \dots, \alpha_k^y, \gamma$. Функцію втрат доцільно приймати у вигляді (3.7), вважаючи її симетричною, тобто $A=B$. Не виникає принципових труднощів, і для узагальнення наведених методів на випадок, коли γ може приймати довільне число значень. Особливого дослідження заслуговує питання вибору сукупності суттєво істотних ознак, за якими повинно проводитись розпізнавання, тобто питання вибору системи функцій $\varphi_i(y)$. Принциповим є той факт, що за умови повноти цієї системи, оптимальне правило розпізнавання, в основних рисах, є незалежним від конкретного вибору зазначеної системи функцій [3,6].

3.3 Статистичні оцінки параметрів сигналів

У цьому підрозділі розглянемо інший, окремий аспект теорії, викладеної у 3.1. Припустимо, що θ являє собою один або декілька $(\theta_1, \dots, \theta_k)$ неперервних параметрів, тобто точку одномірного або багатовимірного (коли $k > 1$) простору. Таким же береться простір значень оцінок θ^* . Оскільки,

прагнуть досягти максимальної близькості між θ^* і θ , функцію вартості слід вибрати залежною лише від різниці аргументів

$$C(\theta^*, \theta) = C(\theta^* - \theta), \quad (3.13)$$

де $C(\theta^*, \theta)$ — функція, що має один мінімум у початку координат. Важливими окремими випадками є квадратична функція вартості

$$C(\theta^*, \theta) = (\theta^* - \theta)^2 \quad (3.14)$$

і проста функція вартості

$$C(\theta^*, \theta) = \text{const} - \delta(\theta^* - \theta), \quad (3.15)$$

які відповідають вимозі до мінімуму середнього квадрата похибки та максимуму апостеріорної густини розподілу ймовірностей оцінюваного параметра, відповідно. Підставляючи (3.14) у формулу мінімального ризику, знаходимо рівняння для оптимальної оцінки

$$\theta^{*2} - 2\theta^* \int_{(\theta)} \theta \mathcal{W}(\theta/y) d\theta + \int_{(\theta)} \theta^2 \mathcal{W}(\theta/y) d\theta = \min_{\theta^*}.$$

Шляхом диференціювання знаходимо, що оптимальна оцінка у цьому випадку збігається з апостеріорним середнім значенням:

$$\theta^* = \theta^*(y) = \int_{(\theta)} \theta \mathcal{W}(\theta/y) d\theta = \mathcal{M}(\theta/y). \quad (3.16)$$

Середній ризик при цьому дорівнює дисперсії похибки

$$\mathcal{R}(\theta^*) = \int_{(\theta)} \int_{(y)} [\theta - \mathcal{M}(\theta/y)]^2 \mathcal{W}(\theta, y) d\theta dy. \quad (3.17)$$

Якщо в рівняння для середнього ризику підставити функцію (3.15), то воно прийме вигляд

$$\mathcal{W}(\theta/y)|_{\theta=\theta^*} \text{ или } \mathcal{W}(\theta)\mathcal{W}(y/\theta)|_{\theta=\theta^*} = \max_{\theta^*}, \quad (3.18)$$

тобто при функції вартості (3.15) оптимальна оцінка відповідає максимуму апостеріорної густини розподілу ймовірностей оцінюваного параметра.

Іноді апріорний розподіл параметра $\mathcal{W}(\theta)$ частково або повністю невідомий. При повній відсутності інформації для обчислення оцінки пропонується у якості апріорного приймати рівномірний розподіл [3], вибір якого не вимагає фіксації яких-небудь функцій або параметрів. При цьому максимізація (3.18) зводиться до максимізації функції правдоподібності

$$\mathcal{W}(y/\theta)|_{\theta=\theta^*} = \max_{\theta^*}. \quad (3.19)$$

У задачах прикладного характеру функція правдоподібності, як правило, має похідні, так що умову екстремальності можна записати також у вигляді

$$\frac{\partial}{\partial \theta} \ln \mathcal{W}(y/\theta)|_{\theta=\theta^*} = 0. \quad (3.20)$$

Це рівняння, корінь якого забезпечує глобальний максимум, прийнято називати рівнянням *максимальної правдоподібності*.

У загальному випадку, оптимальна середньоквадратична оцінка (3.16) відрізняється від оцінок по максимуму апостеріорної ймовірності та максимуму правдоподібності. Однак у деяких випадках, а саме, коли $\mathcal{W}(\mathbf{y}/\theta)$ і $\mathcal{W}(\theta/\mathbf{y})$ – симетричні щодо деякої точки, то всі три критерії приводять до однієї і тієї ж оцінки θ^* . Більше того, до тієї ж оцінки призводить будь-яка інша симетрична відносно початку координат функція вартості, що входить у (3.13). Якщо умова симетрії $\mathcal{W}(\mathbf{y}/\theta)$ виконана не точно, то у багатьох практично важливих випадках, вона виявляється виконаною приблизно. Коли рівень завад відносно малий або коли спостереження проводяться досить довгий час, функцію правдоподібності можна апроксимувати гаусовою кривою з достатньою для практичних потреб точністю

$$\mathcal{W}(\mathbf{y}/\theta) \approx \mathcal{W}(\mathbf{y}/\theta_{\text{МП}}^*) \exp\left\{\frac{1}{2} \frac{\partial^2}{\partial \theta^2} \ln \mathcal{W}(\mathbf{y}/\theta_{\text{МП}}^*) (\theta - \theta_{\text{МП}}^*)^2\right\}, \quad (3.21)$$

де $\theta_{\text{МП}}^* = \theta_{\text{МП}}^*(\mathbf{y})$ — оцінка по максимуму правдоподібності. Пояснимо цю можливість на конкретному прикладі.

Приклад 1. Нехай задана послідовність $\{y_1, \dots, y_n\}$ незалежних спостережень, так що

$$\ln \mathcal{W}(\mathbf{y}/\theta) = \sum_{i=1}^n \ln \mathcal{W}_i(y_i/\theta).$$

Розвинемо цю функцію у ряд Тейлора навколо точки максимуму функції правдоподібності $\theta_{\text{МП}}^*$, тоді будемо мати

$$\begin{aligned} \ln \mathcal{W}(\mathbf{y}/\theta) = \ln \mathcal{W}(\mathbf{y}/\theta_{\text{МП}}^*) + \frac{1}{2} \sum_{i=1}^n \frac{1}{2} \frac{\partial^2}{\partial \theta^2} \ln \mathcal{W}_i(y_i/\theta_{\text{МП}}^*) (\theta - \theta_{\text{МП}}^*)^2 + \\ + \frac{1}{6} \sum_{i=1}^n \frac{\partial^3}{\partial \theta^3} \ln \mathcal{W}_i(y_i/\theta_{\text{МП}}^*) (\theta - \theta_{\text{МП}}^*)^3 + \dots \end{aligned} \quad (3.22)$$

Отже, можна записати

$$\frac{\mathcal{W}(\mathbf{y}/\theta)}{\mathcal{W}(\mathbf{y}/\theta_{\text{МП}}^*)} = \exp\left\{\frac{1}{2} A_2(n) (\theta - \theta_{\text{МП}}^*)^2\right\} \times \exp\left\{\frac{1}{6} A_3(n) (\theta - \theta_{\text{МП}}^*)^3\right\} \dots, \quad (3.23)$$

де

$$A_2(n) = - \sum_{i=1}^n \frac{\partial^2}{\partial \theta^2} \ln \mathcal{W}_i(y_i/\theta_{\text{МП}}^*),$$

$$A_3(n) = - \sum_{i=1}^n \frac{\partial^3}{\partial \theta^3} \ln \mathcal{W}_i(y_i/\theta_{\text{МП}}^*), \quad (3.24)$$

асимптотично лінійно зростають із ростом n .

Перший співмножник $\exp\left\{-\frac{1}{2} A_2(n) (\theta - \theta_{\text{МП}}^*)^2\right\}$ у правій частині (3.23) помітно відрізняється від нуля лише у області, де

$$(\theta^* - \theta) \lesssim \frac{1}{\sqrt{A_2(n)}} \sim \frac{1}{\sqrt{n}},$$

але у цій же області наступний член розкладання задовольняє нерівності

$$|A_3(n)(\theta - \theta_{\text{МП}}^*)^3| \lesssim \left| \frac{A_3(n)}{A^{3/2}_3(n)} \right| \sim \frac{1}{\sqrt{n}} \quad (3.25)$$

По тій же причині при великих n стають нехтовно малими й більш високого порядку члени розкладання. Якщо їх відкинути, то одержимо формулу (3.21).

Аналогічно, можна аргументувати заміну функції правдоподібності гаусовою кривою (3.21), і у випадку, коли $\mathbf{y}(t)$ являє собою реалізацію процесу неперервного у часі процесу. Для припустимості цієї заміни важливо, щоб вирази

$$\frac{-\frac{\partial^k}{\partial \theta^k} \ln \mathcal{W}(\mathbf{y}/\theta_{\text{МП}}^*)}{\left[-\frac{\partial^2}{\partial \theta^2} \ln \mathcal{W}(\mathbf{y}/\theta_{\text{МП}}^*)\right]^{k/2}}, \quad k = 3, 4, \dots$$

були у середньому малими величинами. Ця умова виконується, коли відношення $t/\tau_{\text{кор}}$ тривалості інтервалу спостереження до часу кореляції процесу $\mathbf{y}(t)$ є значно більшим одиниці. Функція правдоподібності має симетрію щодо істинного значення параметра θ . При досить великих n майже будь-який апріорний розподіл буде відносно широким, тобто постійним у області, де помітно відрізняється від нуля функція правдоподібності $\mathcal{W}(\mathbf{y}/\theta)$. Цілком логічно, що за таких умов оцінка за максимумом правдоподібності та будь-яка інша байєсова оцінка із симетричною функцією вартості будуть мало відрізнятися від апостеріорного середнього $\mathcal{M}(\theta/\mathbf{y})$. У цьому випадку, має місце широко відома незалежність оптимальних оцінок від апріорного розподілу та від функції вартості [2,6]. Це дозволяє при виборі того або іншого розподілу та критерію керуватися значною мірою міркуваннями простоти та зручності.

Розглянемо більш детально випадок, коли апріорний розподіл $\mathcal{W}(\theta)$ відсутній і байєсовими критеріями користуватись стає неможливо. У цьому разі, якість оцінювання доцільно характеризувати умовним ризиком. При квадратичній функції вартості отримаємо

$$\mathcal{R}(\theta^*, \theta) = \int_{(\mathbf{y})} [\theta^*(\mathbf{y}) - \theta]^2 \mathcal{W}(\mathbf{y}/\theta) d\mathbf{y}. \quad (3.26)$$

Оцінка $\theta^*(\mathbf{y})$ називається незміщеною, якщо її математичне сподівання співпадає з дійсним значенням оцінюваного параметра, тобто при будь-якому θ повинна виконуватись умова

$$\int_{(\mathbf{y})} \theta^*(\mathbf{y}) \mathcal{W}(\mathbf{y}/\theta) d\mathbf{y} = \theta.$$

Для таких оцінок умовний ризик збігається з дисперсією оцінки

$$\mathcal{R}(\theta^*, \theta) = \int_{(y)} [\theta^*(y) - \int_{(y)} \theta^*(y) \mathcal{W}(y/\theta) dy]^2 \mathcal{W}(y/\theta) dy, \quad (3.27)$$

яка є важливою характеристикою оцінки.

Можна показати, що дисперсія (3.27) будь-якої незміщеної оцінки задовольняє нерівності Крамера–Рао [3,6]:

$$\mathcal{R}(\theta^*, \theta) \geq \frac{1}{\mathcal{F}(\theta)},$$

де

$$\mathcal{F}(\theta) = - \int_{(y)} \left[\frac{\partial^2}{\partial \theta^2} \ln \mathcal{W}(y/\theta) \right] \mathcal{W}(y/\theta) dy \quad (3.28)$$

— так звана, інформаційна міра Фішера. Для найкращої оцінки, яку називають ефективною, виконується рівність

$$\mathcal{R}(\theta^*, \theta) = \frac{1}{\mathcal{F}(\theta)}. \quad (3.29)$$

Ефективна оцінка існує не завжди, а дослідження факту, що оцінка є ефективною часто являє собою складну задачу математичного характеру. Проте, якщо така оцінка все-таки існує, то вона може бути знайдена розв'язком рівняння правдоподібності. Середнє значення другої похідної у виразі (3.28), є не що інше, як коефіцієнт, що визначає кривизну параболи в (3.21). При рівномірному апіорному розподілі θ параболи в (3.21) відповідає гаусів апостеріорний розподіл параметра з дисперсією

$$\sigma = \left\{ \mathcal{M} \left[\frac{\partial^2}{\partial \theta^2} \ln \mathcal{W}(y/\theta) \right] \Big|_{\theta=\theta_{\text{МП}}} \right\}^{-1}. \quad (3.30)$$

Ця формула свідчить про те, що друга похідна від логарифма функції правдоподібності характеризує апостеріорну похибку. У відповідності до (3.26), (3.28) середньоквадратичний розкид координати максимуму правдоподібності щодо дійсного значення параметра θ дорівнює середній ширині піка правдоподібності. Різниця лише у додатковому осередненні другої похідної. У багатовимірному випадку похідна $\frac{\partial^2}{\partial \theta^2} \ln \mathcal{W}(y/\theta)$ замінюється на матрицю других похідних, які після осереднення утворюють інформаційну матрицю Фішера [2,3,6] з елементами

$$I_{ij} = - \int_{(y)} \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \mathcal{W}(y/\theta_1, \dots, \theta_k) \right] \mathcal{W}(y/\theta_1, \dots, \theta_k) dy. \quad (3.31)$$

Найкращі (ефективні) оцінки θ^* задовольняють рівності

$$\|\mathcal{M}\{[\theta_i^*(y) - \mathcal{M}[\theta_i^*(y)/\theta_1, \dots, \theta_k]][\theta_j^*(y) - \mathcal{M}[\theta_j^*(y)/\theta_1, \dots, \theta_k]]\}\| = \|I_{ij}\|^{-1}, \quad (3.32)$$

і за умови їх існування можуть бути знайдені спільним розв'язком рівнянь правдоподібності

$$\frac{\partial}{\partial \theta_k} \ln \mathcal{W}(\mathbf{y}/\theta_1, \dots, \theta_k) = 0, \quad i = 1, \dots, k. \quad (3.33)$$

Співвідношення (3.32) характеризує якість оцінок $\theta_i^*(y)$ і вони тим точніші, чим менші елементи матриці, що розташована у правій частині цього виразу. Щодо збільшення числа вимірювань n , то оцінки зазвичай, мають деякі асимптотичні властивості, до яких відносять асимптотичну незміщеність та ефективність.

Приклад 2. Припустимо, що вимірюються параметри θ сигналу $S(t, \theta)$, прийнятого на фоні білого шуму. Тоді при скалярному θ мінімальна дисперсія згідно (3.30) дорівнює

$$\sigma = \frac{N_0}{b(\theta)}, \quad b(\theta) = \left. \frac{\partial^2 \psi(\theta_1, \theta_2)}{\partial \theta_1 \partial \theta_2} \right|_{\theta_1 = \theta_2 = \theta},$$

де $\psi(\theta_1, \theta_2)$ —функція невизначеності (автокореляції) сигналу. Зокрема, коли $\psi(\theta_1, \theta_2) = \psi(\theta_1 - \theta_2)$, величина b від θ не залежить. Точність тем вища, чим гостріша вершина функції невизначеності. Також, як і в інших прикладах, у випадку білого шуму інформаційна матриця Фішера дорівнює другим похідним від функції невизначеності, що поділені на величину спектральної потужності білого шуму N_0 .

Достатні статистики. Коли процес спостереження $\mathbf{y}(t)$ складається з великої кількості випадкових величин (є процесом на інтервалі великої тривалості), у теорії оцінок надзвичайно велику роль відіграє поняття достатньої статистики [17]. Це такі функції від $\mathbf{y}(t)$ $X_i(\mathbf{y})$, $i = 1, \dots, n$, за допомогою яких функція правдоподібності може бути представлена у вигляді

$$\mathcal{W}(\mathbf{y}/\theta) = C(\mathbf{y})F[X_1(\mathbf{y}), \dots, X_n(\mathbf{y}); \theta], \quad (3.34)$$

де C не залежить від θ , а співмножник F залежить від $\mathbf{y}(t)$ тільки за посередництва $X_1(\mathbf{y}), \dots, X_n(\mathbf{y})$. Якщо $\mathbf{y}(t)$ є процесом неперервним у часі, то $X_1(\mathbf{y}), \dots, X_n(\mathbf{y})$, і $C(\mathbf{y})$ є функціоналами від \mathbf{y} , але F , і у цьому випадку є функцією скінченного числа змінних.

У байєсовій теорії оцінювання від достатніх статистик вимагається, щоб в аналогічній формі виражався сукупний закон розподілу

$$\mathcal{W}(\mathbf{y}, \theta) = C(\mathbf{y})F[X_1(\mathbf{y}), \dots, X_n(\mathbf{y}); \theta].$$

Тоді, очевидно, що апостеріорний розподіл теж буде залежати від $\mathbf{y}(t)$ тільки за посередництва $X_1(\mathbf{y}), \dots, X_n(\mathbf{y})$:

$$\mathbf{w}(\boldsymbol{\theta}/\mathbf{y}) = \frac{F[X_1(\mathbf{y}), \dots, X_n(\mathbf{y}); \boldsymbol{\theta}]}{\int_{(\boldsymbol{\theta})} F[X_1(\mathbf{y}), \dots, X_n(\mathbf{y}); \boldsymbol{\theta}] d\boldsymbol{\theta}}.$$

Отже, функції $X_i(\mathbf{y})$ будуть достатніми для обчислення апостеріорних характеристик, що коротко можна записати у вигляді

$$\mathbf{w}(\boldsymbol{\theta}/\mathbf{y}) = \mathbf{w}[\boldsymbol{\theta}/X(\mathbf{y})]. \quad (3.35)$$

Підставляючи (3.35) у рівняння для середнього ризику, легко побачити, що оптимальна процедура оцінювання залежить від $\mathbf{y}(t)$ тільки через достатні статистики [3]:

$$\boldsymbol{\theta}^*(\mathbf{y}) = \boldsymbol{\theta}^*[X(\mathbf{y})].$$

В небайєсовій теорії також можна обмежитися розглядом процедур типу $\boldsymbol{\theta}^*[X(\mathbf{y})]$. Оскільки число достатніх статистик, як правило, значно менше числа випадкових величин, що входять до складу \mathbf{y} , відшукувати функцію $\boldsymbol{\theta}^*[X(\mathbf{y})]$ значно простіше, ніж безпосередньо функцію $\boldsymbol{\theta}^*(\mathbf{y})$. Тому достатніми статистиками дуже зручно користуватися, якщо вони існують.

Приклад 3. Нехай θ є невідоме середнє значення гаусового розподілу із заданою дисперсією σ^2 . При n незалежних вибірках із цього розподілу функція правдоподібності дорівнює

$$\begin{aligned} \mathcal{W} &= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \theta)^2 \right\} = \\ &= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n y_i^2 \right\} \exp \left\{ \frac{2\theta X(\mathbf{y}) - \theta^2 n}{2\sigma^2} \right\}. \end{aligned} \quad (3.36)$$

У виразі (3.36) $X(\mathbf{y}) = \sum_{i=1}^n y_i^2$, від θ залежить лише друга експонента. Вираз (3.36) являє собою окремий випадок формули (3.34), причому $X(\mathbf{y})$ виконує роль достатньої статистики. Оцінка за максимумом правдоподібності дорівнює $\frac{X(\mathbf{y})}{n}$, тобто дійсно має вигляд $\boldsymbol{\theta}^*[X(\mathbf{y})]$.

3.4 Розрізнення сигналів та синтез сигналів оптимальної форми

Існує декілька підходів до розв'язання задачі розрізнення сигналів [2, 3, 6]. Наприклад, один з них базується на теорії прийняття статистичних рішень і трактує розрізнення як спосіб зіставлення альтернатив про наявність або відсутність даного сигналу за наявності інших сигналів з близькими значеннями параметра θ . Фактично розв'язується задача виявлення в ускладнених умовах, а саме: наявність суміжних завад, за структурою близьких до сигналу, що виявляється. Така постановка задачі приводить до порівняння з порогом коефіцієнта правдоподібності:

$$\Lambda = \frac{\mathcal{W}(\mathbf{y}/\theta_1, \theta_2)}{\mathcal{W}(\mathbf{y}/\theta_2)}, \quad (3.37)$$

де $\mathcal{W}(y/\theta_1, \theta_2)$ — умовна густина розподілу ймовірностей спостережуваного вхідного процесу $y(t) = S_1(t) + S_2(t) + n(t)$ за наявності корисного сигналу $S_1(t)$ та суміжного з ним сигналу $S_2(t)$, що заважає і шуму $n(t)$; $\mathcal{W}(y/\theta_2)$ — відповідна густина розподілу за наявності сигналу, що лише заважає $S_2(t)$ та шуму $n(t)$.

Другий підхід до проблеми розрізнення сигналів опирається на відомі методи ортогоналізації функцій. Згідно [3] оптимальний приймач детермінованого сигналу на фоні шуму є кореляційним. Розглянемо два сигнали $S_1(t)$ і $S_2(t)$. Побудувавши кореляційний приймач для першого з них, у загальному випадку одержимо ненульовий сигнал на виході приймача і при дії другого сигналу $S_2(t)$, якщо

$$\int_0^T S_1(t)S_2(t) dt \neq 0.$$

Проте, «опорну напругу» $S_1(t)$ у приймачі можна замінити на

$$S_1(t) - S_2(t) \sqrt{\rho_{12}} \psi_{12},$$

де

$$\rho_{12} = \int_0^T S_1^2(t) dt / \int_0^T S_2^2(t) dt$$

– відношення енергії сигналів;

$$\psi_{12} = \int_0^T S_1(t)S_2(t) dt / \sqrt{\int_0^T S_1^2(t) dt \int_0^T S_2^2(t) dt}$$

– нормований коефіцієнт взаємної кореляції двох сигналів.

При повній тотожності сигналів $S_1(t) \equiv S_2(t)$, або їх пропорційності ψ_{12} згортається в одиницю, при повній відмінності сигналів або істотному взаємному часовому зсуві $\psi_{12}=0$. Тоді сигнал $S_1(t)$ пройде (правда в дещо ослабленому вигляді) через новий приймач, а сигнал $S_2(t)$, буде повністю пригнічений.

Другий підхід неважко узагальнити й на ситуацію з декількома сигналами $S_i(t)$, $i = 1, \dots, n$ у шумах. При цьому передбачається, що оператор розрізнення містить лінійну частину, що виділяє сигнал $S_0(t)$ і частину, яка виконує процедуру режекції відносно сигналів $S_i(t)$, $i = 1, \dots, n$. Оператор лінійної частини підбирається так, щоб створити максимум відношення сигнал/шум при повній режекції усіх суміжних сигналів. Сигнали можуть бути флуктуаційними та високочастотними. При цьому $S_i(t)$, $i = 1, \dots, n$ за-

мінються на коефіцієнти модуляції $U(t, \theta_i)$, а ступінь схожості сигналів описують через функцію невизначеності.

Незважаючи на відмінність у постановці задачі, обидва зазначені методи синтезу приводять до близьких результатів. В обох випадках приймач містить елементи пригнічення суміжних сигналів. Але якщо з позиції теорії статистичних рішень пригнічення здійснюється до оптимального рівня, що забезпечує, наприклад, найвищу ймовірність правильного виявлення при фіксованому рівні фальшивої тривоги, то при ортогоналізації режекція здійснюється повністю, що трохи погіршує якість приймання при $\psi(\theta_0, \theta_i) \approx 1$. Тим самим другий підхід не враховує наявності шуму, теоретично менш послідовний і гранично досяжної якості обробки не забезпечує. Проте, з іншої сторони, для користування ним не потрібно знати багатомірних розподілів миттєвих значень сигналів.

Приклад 1. Припустимо, що флюктуючий сигнал з коефіцієнтом модуляції $U(t, \theta_1)$, функцією невизначеності $\psi_{12}(\theta_1 - \theta_2)$ і відношенням сигнал/шум ρ_1 виділяється на фоні подібного сигналу з характеристиками $U(t, \theta_2)$, ρ_2 і шуму. Розподіл ймовірностей усіх компонентів вважаємо нормальним. Через наявності суміжного сигналу ефективне відношення сигнал/сумарна завада при виділенні першого сигналу виявляється меншим у порівнянні з випадком наявності лише шуму до величини

$$\rho_{\text{эф}0} = \rho - \frac{\rho_2}{1 + \rho_2} |\psi_{12}|^2$$

при оптимальному приймачі, що здійснює процедуру розрізнення і до величини

$$\rho_{\text{эф}p} = \rho_1 [1 - |\psi_{12}|^2],$$

при приймачі, що забезпечує повну режекцію.

Результати як попередніх підрозділів, так і даного показують, що якісні показники процедур оцінювання та розрізнення багато у чому залежать від форми сигналів через функцію невизначеності. Так, найкращим з погляду процесу вимірювань за наявності шумів, виявляється сигнал з найбільш гострою вершиною функції невизначеності, а з погляду розрізнення, крім того, ще й з найменшими бічними максимумами (викидами) цієї функції. Це дозволяє ставити задачу про вибір сигналу оптимальної форми, тобто крім оптимізації приймачів оптимізувати ще й передавальний тракт радіоелектронних систем.

Дійсно, для підвищення точності вимірювань потрібно зажадати

$$-\frac{\partial^2 \psi(\Delta)}{\partial \Delta_i \partial \Delta_j} = \max_{\psi}; \quad i, j = 1, \dots, n, \quad (3.38)$$

а для поліпшення роздільної здатності можна, наприклад, додатково зажадати найменшого значення найбільшого з побічних максимумів:

$$\max_{\Delta \neq 0} |\psi(\Delta)|^2 = \min_{\psi}. \quad (3.39)$$

Звідси отримуємо необхідний вигляд функції невизначеності, а по ній стане можливим визначати і сам сигнал. Проте, відразу ж слід зауважити, що загального та однозначного методу відновлення сигналу за його функцією невизначеності не існує. Більше того, відомо, що близьку за виглядом функцію невизначеності можуть мати зовсім різні сигнали. Тому, як правило, обмежуються оптимізацією сигналів деяких окремих типів (із частотною, фазо-кодуючою та ін. модуляціями), домагаючись виконання (3.39). Тут відома велика кількість результатів, розкиданих по різних науково-технічних журналах та літературним джерелам [11].

Іноді при синтезі сигналів трапляються обмеження принципового характеру. Наприклад, нехай двома параметрами сигналу є часова затримка τ та частотний зсув Ω :

$$U(t, \tau, \Omega) = U_0(t - \tau)e^{j\Omega t}, \quad (3.40)$$

де функція $U_0(t)$ вважається повністю відомою. Відповідно до цього випадку, функція невизначеності сигналів $U(t, \tau, \Omega)$ і $U(t, \tau + \Delta_\tau, \Omega + \Delta_\Omega)$ дорівнює

$$\psi(\Delta_\tau, \Delta_\Omega) = \int_{-\infty}^{\infty} U_0(t)U_0^*(t + \Delta_\tau)e^{j\Delta_\Omega t} dt,$$

У роботі [6], доведена властивість цієї функції щодо збереження об'єму

$$\int_{-\infty}^{\infty} |\psi(\Delta_\tau, \Delta_\Omega)|^2 d\Delta_\tau, d\Delta_\Omega = \text{const.}$$

Тим самим, при синтезі, що опирається на співвідношення (3.38)–(3.39) стає неможливим стиснення ψ одразу в усіх напрямках. Можна говорити лише, про збереження принципової невизначеності при одночасних вимірюваннях часової затримки та частотного зсуву, що й пояснює введення даного терміна.

3.5 Лінійна оптимальна фільтрація

У теорії лінійної оптимальної фільтрації заздалегідь обмежуються розглядом лише лінійних перетворень $u = \gamma(y)$. Зокрема, коли мають справу зі значеннями випадкової функції $y(t)$ на інтервалі (a,b) , u – скалярна випадкова величина, то зазначене перетворення має вигляд

$$u = \gamma(y) = \int_a^b h(t) y(t) dt, \quad (3.41)$$

де $h(t)$ – деяка функція, яку необхідно знайти з міркувань оптимальності.

Якщо розглядається не скалярна оцінка u , а функція $u(t)$, то формулу (3.41) слід застосовувати при фіксованому значенні $t = t_0$, тобто це буде означати, що

$$u(t) = \int_a^b h(t_0, t) y(t) dt,$$

де $h(t_0, t)$ – імпульсна характеристика лінійної системи.

Розглянемо оптимальне перетворення, що відповідає середньоквадратичному критерію, тобто функції втрат (3.32). Тепер уже не можна користуватися формулою (3.15), тому що перетворення, яке безпосередньо з неї випливає, може виявитися нелінійним, а слід мінімізувати ризик (3.2) по безлічі лінійних перетворень (3.41).

Підставимо (3.14) і (3.41) в (3.2), тоді отримаємо

$$\begin{aligned} \mathcal{R}(\gamma) &= \mathcal{M}\{\theta^2\} - 2 \int_a^b h(t) \mathcal{M}\{[y(t), \theta]\} dt + \\ &+ \int_a^b \int_a^b h(t) h(t') \mathcal{M}\{y(t) y(t')\} dt dt'. \end{aligned} \quad (3.42)$$

У результаті мінімізації (3.42) по усіх можливих функціях $h(t)$ одержується, так зване рівняння регресії

$$\int_a^b \mathcal{M}\{y(t) y(t')\} h(t') dt dt' = \mathcal{M}\{y(t), \theta\}, \quad (3.43)$$

з якого і визначається оптимальна функція $h(t)$. Середній ризик (3.42) для оптимального перетворення становить

$$\mathcal{R}(\gamma) = \mathcal{M}\{\theta^2\} - \int_a^b h(t) \mathcal{M}\{[y(t), \theta]\} dt. \quad (3.44)$$

Корисно зазначити, що для знаходження оптимального лінійного перетворення виявилось достатнім знання лише других моментів вхідного процесу, а не їх багатомірних розподілів.

Дискретний аналог рівняння (3.43) вдається безпосередньо розв'язати, якщо $y(t)$ являє собою набір скінченного числа випадкових величин. При нескінченнім числі величин та у розглянутому вище континуальному (неперервному) випадку його розв'язок ускладнюється. Приведемо окремі випадки такого розв'язку.

1. Інтервал спостереження нескінченний: $a = -\infty, b = \infty$, процес $y(t)$ стаціонарний, так що

$$\mathcal{M}\{y(t)y(t')\} = R_{yy}(t - t').$$

Потрібно оцінити значення (у деякий момент часу t_0) процесу $\theta(t)$, стаціонарно пов'язаного з $y(t)$, так що

$$\theta = \theta(t_0), \quad \mathcal{M}\{y(t), \theta\} = \mathcal{M}\{y(t), \theta(t_0)\} = R_{y\theta}(t - t_0).$$

Тоді рівняння (3.43) можна розв'язати, використовуючи апарат перетворень Фур'є [1]. Функція $h(t) = h_0(t_0 - t)$ виявляється залежною лише від різниці $t_0 - t = \tau$ і відповідна їй передатна функція лінійної системи дорівнює

$$H(j\omega) = \int_{-\infty}^{\infty} e^{j\omega\tau} h_0(\tau) d\tau = \frac{S_{\theta y}(\omega)}{S_{yy}(\omega)}, \quad (3.45)$$

де

$$S_{yy}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_{yy}(\tau) d\tau;$$

$$S_{\theta y}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_{\theta y}(\tau) d\tau. \quad (3.46)$$

Для середніх втрат згідно (3.44) маємо формулу

$$\mathcal{R}(\gamma) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [S_{\theta\theta}(\omega) - |S_{\theta y}(\omega)|^2 / S_{yy}(\omega)] d\omega. \quad (3.47)$$

Така фільтрація у теорії прийняття статистичних рішень називається лінійною фільтрацією без урахування умови фізичної реалізації, тому що для оцінки значення θ використовуються не тільки минулі $t < t_0$ стосовно моменту часу t_0 , але й майбутні $t > t_0$ значення спостережуваного сигналу. Отже, щоб зробити таку оцінку, потрібно мати у своєму розпорядженні запис усього процесу спостереження $y(t)$.

2. Якщо спостережуваними величинами $y(t)$ (до моменту t_0 можна вважати лише значення $y(t)$ у попередні моменти часу $t < t_0$, те це відповідає умові урахування фізичної реалізації. Тут як і раніше беремо $\theta = \theta(t_0)$, $a = -\infty$, але $b = t_0$. Для стаціонарних процесів $\theta(t)$, $y(t)$, відповідна до цього випадку, модифікація рівняння (3.43) називається рівнянням Вінера-Хопфа:

$$\int_0^{\infty} R_{yy}(\tau - \tau') h_0(\tau') d\tau' = R_{\theta y}(\tau), \quad \tau > 0. \quad (3.48)$$

У якості його розв'язку, замість (3.47), маємо більш складну формулу

$$H = \frac{1}{\psi(j\omega)} \left[\frac{S_{\theta y}(\omega)}{\psi(-j\omega)} \right]_+ , \quad (3.49)$$

Де $\psi(j\omega)\psi(-j\omega) = S_{yy}(\omega)$ – спеціальне розвинення (факторизація) на співмножники з нулями та полюсами у верхній (перший співмножник) і в нижній (другий співмножник) півплощини комплексної змінної ω . Дужки $[\dots]_+$ в (3.49) означають, що беруться ті складові розкладу на найпростіші дроби, які мають полюси у верхній півплощині. Факторизацію та операцію $[\dots]_+$ можна також записувати у вигляді інтегралу. Однак у випадку раціональних спектральних функцій $S_{yy}(\omega)$, $S_{\theta y}(\omega)$ обчислення зручніше проводити, минаючи інтегрування, шляхом розкладання функції $S_{yy}(\omega)$ на елементарні співмножники, а функції $S_{\theta y}(\omega)/\psi(-j\omega)$ на найпростіші дроби. Наведений розв'язок був отриманий Вінером [2]. Відповідний дискретний аналог розв'язку (для стаціонарних послідовностей) був знайдений Колмогоровим [3].

3. У випадку скінченного інтервалу спостереження $[a, b]$ розв'язок рівняння (3.48) може бути отриманий, якщо $\theta(t)$, $y(t)$ є стаціонарними процесами з раціональною спектральною густиною. У випадку нестаціонарних процесів точний розв'язок можна знайти методами теорії умовних процесів, коли процеси $\theta(t)$, $y(t)$ зводяться до багатовимірного процесу Маркова. Запропоновані ще й інші точні або наближені методи розв'язку для низки окремих випадків [6, 17]. Усе вищезазначене безпосередньо поширюється і на випадок багатомірних процесів $\theta(t)$, $y(t)$ та на випадок процесів у дискретному часі (випадкових послідовностей). Теорія лінійної фільтрації у радіотехнічних системах застосовується звичайно у тих випадках, коли завада є адитивною і незалежною, тобто спостережуваний процес є сумою $y(t) = \theta(t) + n(t)$ корисного сигналу $\theta(t)$ і статистично незалежної від нього завади — шуму $n(t)$. У силу незалежності маємо при нульових середніх значеннях

$$\mathcal{M}\{y(t)y(t')\} = R_{\theta\theta}(t, t') + R_{nn}(t, t'),$$

$$\mathcal{M}\{y(t)\theta(t')\} = R_{\theta\theta}(t, t'),$$

де $R_{\theta\theta}(t, t')$, $R_{nn}(t, t')$ — кореляційні функції $\theta(t)$ і $n(t)$, відповідно.

Рівняння регресії (3.43) при $\theta = \theta(t_0)$ приймає вигляд

$$\int_a^b [R_{\theta\theta}(t, t') + R_{nn}(t, t')] h_0(t_0, t) dt' = R_{\theta\theta}(t, t_0), \quad (3.50)$$

у окремому випадку може бути $b = t_0$.

Важливо відзначити, що у низці випадків оптимальне перетворення $\theta^* = \theta^*(y)$, що знаходиться із більш загального співвідношення (3.5), виявляється лінійним, отже теорія лінійної фільтрації приводить до оптимального перетворення. Це має місце, коли спільний розподіл процесів $\theta(t), y(t)$ гауссовий, а функція вартості є квадратичною.

3.6 Нелінійна оптимальна фільтрація при гаусовій апроксимації

У загальному випадку оптимальне байєсове перетворення, що описується формулою (3.5) є нелінійним. Якщо, у цьому разі обмежитись лише класом лінійних перетворень, то це призведе до втрати оптимальності. Окрім того, у низці випадків заздальгідь використовувати лише лінійні перетворення буває недоцільним або практично неприпустимо. Нелінійне оптимальне перетворення при спостереженні мінливих у часі випадкових процесів за умови дії шумів повинно відшукуватись методами оптимальної нелінійної фільтрації. Існує декілька підходів до проблеми нелінійної оптимальної фільтрації. Можна, наприклад, заздальгідь обмежитись деяким класом нелінійних перетворень і відшукувати найкращі з них у заданому наперед класі [6]. При цьому отримують рівняння, які важко розв'язуються з математичної точки зору і залишається незрозумілим, наскільки оптимальне перетворення у даному класі близьке до оптимального перетворення, яке знаходиться без зазначених обмежень.

При інших підходах не вводяться обмеження на клас перетворень, але використовуються марковські властивості спостережуваних процесів, або накладається умова високої апостеріорної точності, що дозволяє застосовувати гаусову апроксимацію багатовимірного апостеріорного розподілу і на певному етапі лінійні методи. При цьому, підсумкове оптимальне перетворення, як видно буде з нижченаведеного, являє собою послідовність, що складається з нелінійного та лінійного перетворень. Дана теорія, яку назвали квазілінійною є, на відміну від марковської теорії, наближеною та асимптотичною, і у неявній формі використовує методи малого або великого параметра. У ролі великого параметра виступає відповідним чином підібране відношення сигнал / шум.

У варіанті дискретного часу вхідний (спостережуваний) $y(t)$ та вихідний (результат фільтрації) $\theta(t)$ процеси являють собою вектори. Нехай потрібно знайти оптимальну оцінку $\theta^*(y_1, \dots, y_n)$ останнього значення θ_n процесу фільтрації. При високій апостеріорній точності (малої похибки вимірювань) конкретний вибір критерію оптимальності стає несуттєвим, і можна обмежитись критерієм середньоквадратичної похибки, який приводить до оптимального оператора у вигляді пошуку апостеріорного середнього. У цьому випадку його можна записати

$$\theta_n(y_1, \dots, y_n) = \frac{\int_n \theta_n \mathbf{w}_n(\theta_1, \dots, \theta_n) \mathbf{w}_n(y_1, \dots, y_n / \theta_1, \dots, \theta_n) d\theta_1, \dots, d\theta_n}{\int_n \mathbf{w}_n(\theta_1, \dots, \theta_n) \mathbf{w}_n(y_1, \dots, y_n / \theta_1, \dots, \theta_n) d\theta_1, \dots, d\theta_n}. \quad (3.51)$$

На величину інтегралів, що входять до (3.51) впливає, в основному, та область n -мірного простору значень $(\theta_1, \dots, \theta_n)$, де підінтегральна функція

$$\mathbf{w}_n(\theta_1, \dots, \theta_n) \mathbf{w}_n(y_1, \dots, y_n / \theta_1, \dots, \theta_n) \quad (3.52)$$

(y_1, \dots, y_n – фіксовані) приймає максимальне значення. Виберемо деяку точку n -мірного простору $\mathbf{v}(v_1, \dots, v_n)$, яка розташована поблизу точки максимуму функції (3.52). Спосіб вибору цієї точки буде пояснений надалі. Поблизу \mathbf{v} є можливість провести апроксимацію логарифма функції (3.52) квадратичною функцією (багатовимірною параболою):

$$\begin{aligned} & \mathbf{w}(\theta_1, \dots, \theta_n) \mathbf{w}_n(y_1, \dots, y_n / \theta_1, \dots, \theta_n) d\theta_1, \dots, d\theta_n \approx \\ & \approx \mathbf{w}(v_1, \dots, v_n) \mathbf{w}_n(y_1, \dots, y_n / v_1, \dots, v_n) \times \\ & \times \exp\left\{ \sum_{i=1}^n (a_i^{(n)} + b_i^{(n)}) (\theta_i - v_i) - \frac{1}{2} \sum_{i,j=1}^n (A_{ij}^{(n)} + B_{ij}^{(n)}) (\theta_i - v_i) (\theta_j - v_j) \right\}, \end{aligned} \quad (3.53)$$

де

$$a_i^{(n)} = \frac{\partial \ln \mathbf{w}_n(v_1, \dots, v_n)}{\partial \theta_i}, \quad b_i^{(n)} = \frac{\partial \ln \mathbf{w}_n(y_1, \dots, y_n / v_1, \dots, v_n)}{\partial \theta_i},$$

$$A_{ij}^{(n)} = -\frac{\partial^2 \ln \mathbf{w}_n(v_1, \dots, v_n)}{\partial \theta_i \partial \theta_j}, \quad B_{ij}^{(n)} = \frac{\partial^2 \ln \mathbf{w}_n(y_1, \dots, y_n / v_1, \dots, v_n)}{\partial \theta_i \partial \theta_j}.$$

Якщо підставити (3.53) у (3.51) та виконати інтегрування, то отримаємо

$$\theta_n^* (y_1, \dots, y_n) = v_n + \sum_{j=1}^n C_{n_j}^{(n)} (a_j^{(n)} + b_j^{(n)}), \quad (3.54)$$

де $C_{n_j}^{(n)}$ — елемент матриці, оберненої до $\|A_{ij}^{(n)} + B_{ij}^{(n)}\|$.

З практичної точки зору, зручно вибирати у якості координат n -мірної точки \mathbf{v} значення оптимальних оцінок, знайдених на усіх кроках, що передували n -му, тобто: $v_i = \theta_{i-1}^*(y_1, \dots, y_{i-1})$, $v_1 = \theta_0^*$, $i = 2, \dots, n$, де θ_0^* — початкове значення, що дорівнює, наприклад, апріорному середньому у початковий момент часу. Співвідношення

$$\theta_n^* = \theta_{n-1}^* + \sum_{j=1}^n C_{n_j}^{(n)} (a_j^{(n)} + b_j^{(n)}), \quad (3.55)$$

при $a_j^{(n)}$, $b_{ij}^{(n)}$, $C_{n_j}^{(n)}$, що визначаються похідними (3.53) у точці $\mathbf{v} = \{\theta_0^*, \dots, \theta_{n-1}^*\}$, і описує процедуру оптимальної нелінійної фільтрації.

Особливо прості результати отримуються із (3.55) у важливому окремому випадку, вивченому І. А. Большаковим та В. Г. Репіним [6]. У цьому разі апіорний розподіл передбачається гаусовим, так що

$$a_k^{(n)} = -\sum_{k=1}^n V_{ik}^{(n)} (\theta_{n-1}^* - m_k), \quad A_{kj}^{(n)} = V_{ij}^{(n)}, \quad (3.56)$$

де $V_{ij}^{(n)}$ ($i, j = 1, \dots, n$) – елемент матриці не залежний від v і обернений до кореляційної матриці $\|R(t_i, t_j)\| = \|R_{ij}\|$ для інформаційного параметра $\theta(t)$; m_i – апіорні середні значення $\theta(t)$. Передбачається, що супутні, неінформативні параметри та завади змінюються значно швидше ніж вимірювана величина $\theta(t)$, так що

$$\mathbf{w}_n(y_1, \dots, y_n / v_1, \dots, v_n) = \prod_{i=1}^n \mathbf{w}_i(y_i / \theta_i), \quad (3.57)$$

$$b_i^{(n)} = b_i, \quad B_{kj}^{(n)} = K_i \delta_{ij},$$

$$\text{де } b_i = \frac{\partial \ln \mathbf{w}_i(y_i / \theta_{i-1}^*)}{\partial \theta_i}, \quad K_i = -\frac{\partial^2 \ln \mathbf{w}_i(y_i / \theta_{i-1}^*)}{\partial \theta_i^2}. \quad (3.58)$$

Якщо підставити (3.56) – (3.57) в (3.55), то остаточно отримуємо

$$\theta_n^* = \sum_{j=1}^n C_{n_j}^{(n)} [b_j + K_j (\theta_{n-1}^* - m_j)] + m_n, \quad (3.59)$$

де $C_{n_j}^{(n)}$ ($j = 1, \dots, n$) задовольняє рівнянню

$$C_{n_j}^{(n)} + \sum_{k=1}^n C_{n_k}^{(n)} K_k R_{kj} = R_{n_j}. \quad (3.60)$$

Знайдений вираз (3.59) описує алгоритм роботи замкненого оптимального фільтра-вимірювача з двома колами зворотного зв'язку (рис. 3.1).

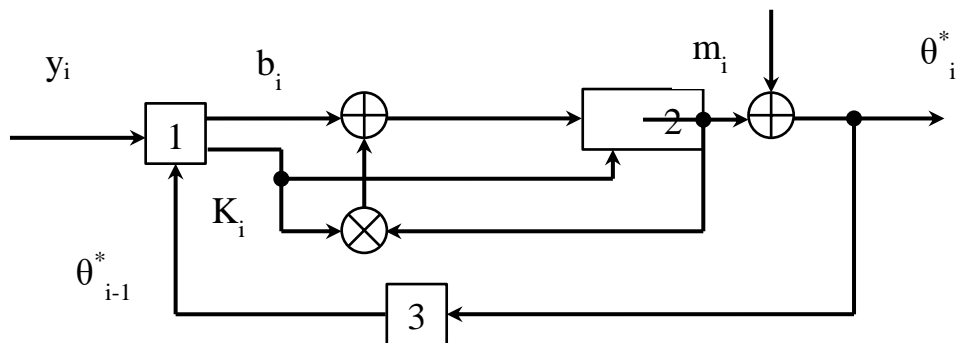


Рисунок 3.1– Структурна схема фільтра – вимірювача з двом колами зворотного зв'язку: 1– оптимальний нелінійний дискримінатор; 2 – ланка згладжування; 3– лінія затримки.

Окрім цього варіанту може бути побудований еквівалентний йому варіант з одним колом зворотного зв'язку. Щоб одержати відповідну йому формулу, потрібно записати рівняння, аналогічні (3.59) не тільки для даного кроку, але й для усіх попередніх кроків $i = 1, \dots, n$ та розв'язати систему рівнянь відносно θ_n^* . Це дає

$$\theta_i^* = \sum_{j=1}^n G_{ij} b_j + m_i, \quad i = 1, \dots, n, \quad (3.61)$$

де G_{ij} задовольняють рівнянню

$$C^{(i)}_{ij} + \sum_{k=j}^n C^{(i)}_{ik} K_k G_{kj} = G_{ij}.$$

Схема фільтра-вимірювача, побудована згідно (3.61) зображена на рис. 3.2.

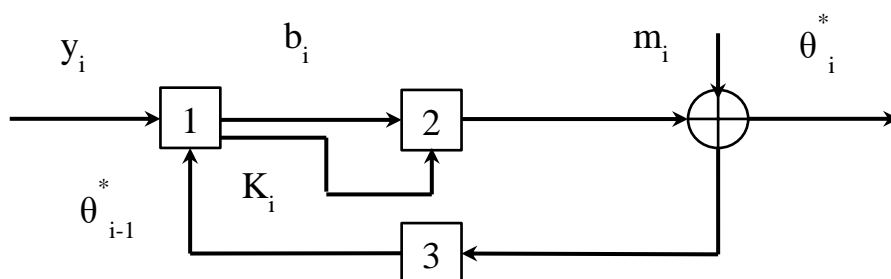


Рисунок 3.2– Структурна схема фільтра-вимірювача з одним колом зворотного зв'язку: 1– оптимальний нелінійний дискримінатор; 2 – ланка згладжування; 3– лінія затримки.

Наведені схеми досить подібні до відомих у техніці і, зокрема, містять оптимальний дискримінатор, який виробляє величину b_i , що у середньому відповідає відхиленню значення θ_{i-1}^* від дійсного значення θ_i . Далі слідом іде лінійне коло згладжування з реакцією G_{ij} , яке скеровується за допомогою функції K_i . Остання є показником точності i -го поодинокого виміру і створюється тим же дискримінатором. Схема завершується суматором, у який вводиться апріорне середнє значення m_i . При невеликих варіаціях поточної точності можна замінити K_i на апріорне середнє значення цієї величини, тоді коло згладжування стає незалежними від вхідного процесу. У ланці 3 здійснюється затримка на період.

При переході до неперервного часу $(t_i - t_{i-1}) = \Delta \rightarrow 0$ вводяться функції $C(t, \tau), G(t, \tau), \theta^*(t), K(t), b(t)$, такі, що

$$\left. \begin{aligned} C(t_i, t_j) &= C_{ij}^{(i)}, \quad G(t_i, t_j) = G_{ij}, \quad \theta^*(t_n) = \theta_n^*, \\ \int_{t_{k-1}}^{t_k} K(t) dt &= K_i, \quad \int_{t_{k-1}}^{t_k} b(t) dt = b_i. \end{aligned} \right\} \quad (3.62)$$

У результаті маємо співвідношення:

$$\left. \begin{aligned} \theta^*(t) &= \int_a^t C(t, \tau) \{ [b(\tau) + K(\tau) [\theta^*(\tau) - m(\tau)] \} dt + m(t), \\ C(t, \tau) + \int_a^t C(t, s) K(s) R(s, \tau) ds &= R(t, \tau), \\ \theta^*(t) &= \int_a^t G(t, \tau) b(\tau) d\tau + m(t), \\ C(t, \tau) + \int_{\tau}^t C(t, s) K(s) G(s, \tau) ds &= G(t, \tau). \end{aligned} \right\} \quad (3.63)$$

Апостеріорна похибка, яка характеризується дисперсією

$$\sigma^2 = \mathcal{M} \left\{ [\theta^*(t) - \theta(t)]^2 \right\},$$

для дискретного та неперервного часу визначається співвідношеннями

$$\sigma_{\text{вих}}^2(n) = \mathcal{M} \left\{ C_{nn}^{(n)}(y) \right\}, \quad \sigma_{\text{вих}}^2(t) = \mathcal{M} \{ C(t, t, y) \},$$

відповідно. Операція осереднення виконується по множині вхідних процесів. Важливо відзначити, що інерційна частина фільтра-вимірювача при гаусовому розподілі θ є лінійною. Більше того, якщо позначити через $h(t, \tau) = C(t, \tau)K(\tau)$, то приходимо від (3.63) до рівняння фільтрації

$$\int_a^t h(t, s) \left[\frac{\delta(s-\tau)}{K(s)} + R(s, \tau) \right] ds = R(t, \tau), \quad (3.64)$$

що співпадає з (3.50), і характеризує виділення сигналу з функцією кореляції $R(t, \tau)$ з нестационарного білого шуму, який має кореляційну функцію $\delta(s - \tau)/K(s)$. Проте, у вимірювачеві присутній дискримінатор (3.58), як правило, нелінійний відносно $y(t)$, так що нелінійним виявляється і вимірювач, у цілому. Слід зауважити, що вищезазначені оцінки $\theta_i^*(y_1, \dots, y_i)$, отримані на i -му кроці, надалі не уточнюються, хоча вони і впливають на вибір наступних оцінок. Точка \mathbf{v} , у якій проводиться розвинення (3.53), повинна перебувати поблизу максимуму виразу, що знаходиться у правій частині (3.53). Оскільки кожна координата θ_i точки максимуму залежить від усіх y_1, \dots, y_n , тому більш обґрунтованим може бути інший вибір точки \mathbf{v} ніж (3.54). Доцільно на кожному n -му кроці знаходити нові оцінки $\theta_i^{*(n)}(y_1, \dots, y_n)$ усіх параметрів $\{\theta_k\}$, які залежать від усіх y_1, \dots, y_n . Саме тому, у якості v_i беруть аналогічні оцінки попереднього кроку, вважаючи що

$$\begin{aligned} v_i &= \theta_i^{*(n-1)}(y_1, \dots, y_{n-1}); \\ &\dots\dots\dots \\ v_n &= \theta_{n-1}^{*(n-1)}(y_1, \dots, y_{n-1}), \quad i = 1, \dots, n-1. \end{aligned} \quad (3.65)$$

За аналогією з (3.54) – (3.55) будемо мати рекурентні перетворення

$$\theta_i^{*(n)}(y_1, \dots, y_n) = v_i + \sum_{j=1}^n C_{ij}^{(n)}(a_j^{(n)} + b_j^{(n)}), \quad (3.66)$$

у яких значення v_k визначаються співвідношеннями (3.65). Таким чином, отримана послідовність оцінок $\theta_i^{*(n)}(y_1, \dots, y_n)$, $n=1,2, \dots$ буде послідовністю більш точних оцінок. Ще більш точні результати отримаємо, якщо підставимо $v_i = \theta_i^{*(n)}$, $i = 1, \dots, n$ в (3.66) та проведемо лінеаризацію даних рівнянь щодо невідомих відхилень $\theta_i^{*(n)} - \theta_i^{*(n-1)}$, після чого отримані рівняння розв'язуються лінійними методами.

3.7 Нелінійна оптимальна фільтрація марковських процесів

Інші ефективні методи пошуку оптимальних нелінійних перетворень засновані на застосуванні теорії умовних марковських процесів, яка була розроблена наприкінці 60–х років минулого століття Р. Л. Стратоновичем. Тут вважається, що процес спостереження $y(t)$ разом з процесом, що підлягає фільтрації $\theta(t)$, є компонентами багатовимірного марковського процесу. Зупинимось на важливому окремому випадку, коли процес $\theta(t)$, сам по собі, є марковським процесом і має густину розподілу ймовірності переходу p , яка допускає опис у неперервному часі рівнянням

$$\frac{\partial p}{\partial t} = Lp(\theta), \quad (3.67)$$

де L – деякий диференціальний оператор. Метод заснований на розв'язку рівняння, якому задовольняють апостеріорні ймовірності

$$w_t(\theta)d\theta = P\{\theta(t) \in \theta, \theta + d\theta / y(\tau), \tau < t\}. \quad (3.68)$$

Вигляд цього рівняння дається теорією умовних марковських процесів. Зокрема, якщо час дискретний, а завада $n(t)$ має у ці ж моменти незалежні відліки, то функція правдоподібності розпадається на добуток

$$P_n(y_1, \dots, y_n / \theta_1, \dots, \theta_n) = \prod_{i=1}^n P_i(y_i / \theta_i), \quad (3.69)$$

тобто зазначене рівняння має вигляд

$$w_{n+1}(\theta) = C_{n+1}P_{n+1}(y_{n+1} / \theta) \int_{-\infty}^{\infty} w_n(\theta')p(\theta', \theta) d\theta', \quad (3.70)$$

де p – апіорна густина розподілу ймовірності переходу за один крок; C_{n+1} — незалежна від θ величина, обумовлена процесом нормування до одиниці $\int_{-\infty}^{\infty} w_{n+1}(\theta)d\theta = 1$. Таким чином, трансформація апостеріорного розподілу (3.70) за один крок визначається апіорними ймовірностями переходу та результатом спостереження y_{n+1} на поточний момент часу,

який надає нову інформацію про значення $\theta(t)$. У тому випадку, коли час є неперервним, рівняння (3.70) замінюється на нелінійне диференціальне рівняння. Коли завада являє собою білий шум (або, у більш загальному випадку, змінюється значно швидше ніж вимірювана величина), це рівняння має вигляд

$$\frac{\partial w_t(\theta)}{\partial t} = Lw_t(\theta) + [F(y, \theta, t) - \int_{-\infty}^{\infty} F(y, \theta', t)w_t(\theta')d\theta']w_t(\theta). \quad (3.71)$$

Тут L - лінійний диференціальний оператор, що входить до виразу (3.67), а F - функція, яка пов'язана з функцією правдоподібності на інтервалі $(t_n, t_n + \Delta)$ співвідношенням

$$\int_{t_n}^{t_n+\Delta} F[y(s), \theta, s]ds = \ln P_{n+1}(y_{n+1}/\theta), \quad \Delta \rightarrow 0, \quad (3.72)$$

де Δ береться більшим ніж інтервал кореляції завади.

Перший доданок $Lw_t(\theta)$ в (3.71) впливає на апостеріорний розподіл, ведучи до його розширення за рахунок збільшення невизначеності через апіорі можливі переходи, другий навпаки сприяє його звуженню (зменшенню невизначеності) за рахунок надходження нової інформації. Рівняння нелінійної фільтрації (3.71) у однакої мірі справедливе як для одномірного, так і багатовимірного марковського процесу. За винятком окремих випадків, розв'язувати рівняння (3.70) або (3.71) аналітично не має потреби. Досить реалізувати радіотехнічну систему, що автоматично здійснює перетворення, які відповідають цим рівнянням. Вказана система являє собою найважливішу ланку оптимального нелінійного фільтра-вимірювача (рис.3.1). Дійсно, оскільки співвідношення (3.5), (3.68) вважаються справедливими за будь-якого довільного критерію $C(\theta, \theta^*)$, то для знаходження оптимального значення оцінки θ^* досить задовольнити вимозі

$$\int_{-\infty}^{\infty} C(\theta_t, \theta_t^*)w_t(\theta_t)d\theta_t = \min_{\theta_t^*}. \quad (3.73)$$

Ланка 2 (рис.3.1), що відшукує мінімум (3.3), автоматично створює таку оцінку.

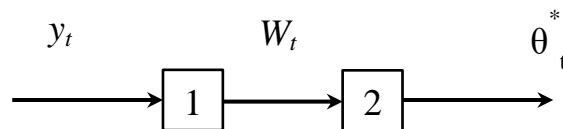


Рисунок 3.3 – Схема створення оптимальної оцінки

Коли $\theta(t)$ є марковським процесом зі скінченним числом станів m , тоді ймовірності (3.68) утворюють скінченне число $(m - 1)$ функцій часу (сигналів) і рівняння (3.70) або (3.71) можуть моделюватися безпосередньо. Якщо марковський процес $\theta(t)$ має нескінченне число станів як, наприклад, дифузійний процес, що описується рівнянням Фоккера-Планка-

Колмогорова, то потрібне попереднє квантування простору станів процесу $\theta(t)$, або перехід від густини розподілу $w_t(\theta)$ до деяких параметрів $\alpha(\alpha_1, \dots, \alpha_n)$, що замінюють цю густину. Наприклад, можна шукати $w_t(\theta)$ у вигляді

$$w_t(\theta) = \text{const} \exp\{\sum_{j=1}^n \alpha_j(t) \varphi_j(\theta)\}, \quad (3.74)$$

де φ_j — задана система функцій, $i=1, \dots, n$

Підставляючи (3.74) у (3.70) або (3.71), можна отримувати точні або наближені рівняння для $\alpha_j(\alpha)$. У варіанті неперервного часу вони будуть мати вигляд

$$\frac{\partial}{\partial t} \alpha_j(t) = \Phi_j(\alpha), \quad j=1, \dots, n, \quad (3.75)$$

де $\Phi_j(\alpha)$ — відомі функції. Система цих рівнянь буде точною, якщо має місце точне розвинення у ряд за обраною системою функцій типу

$$\exp\{-\sum_j \alpha_j \varphi_j(\theta)\} L \exp\{\sum_j \alpha_j \varphi_j(\theta)\} + F\{y, \theta\} = \sum_j \Phi_j \varphi_j(\theta) + \text{const} \quad (3.76)$$

де Φ_j — деякі коефіцієнти. У цьому разі рівняння (3.75) отримують шляхом підстановки (3.74) у рівняння (3.71) поділене на w_t та наступним прирівнюванням коефіцієнтів при різних функціях φ_j .

У низці випадків, коли апріорні переходи відсутні ($L=0$), для переходу до скінченного числа точних рівнянь (3.75) досить виконати розвинення у ряд (3.76) тільки функції $F(y, \theta)$ по скінченній системі функцій $\varphi_j(\theta)$. Якщо розвинення у ряд (3.76) здійснити не вдається, то шляхом вибору достатнього числа функцій $\varphi_j(\theta)$, можна добитися, як завгодно малої похибки при переході від точного рівняння (3.71) до наближених (3.75). У тому разі коли оператор L відповідає дифузійному процесові, де B не залежить від θ , A — лінійна функція θ , а функція P є поліномом від θ ступеня не вище другого, то у формулі (3.74) потрібно вибрати функції $\varphi_1 = \theta$, $\varphi_2 = \theta^2$, що відповідає гауссовому закону розподілу ймовірностей. При цьому отримуються точні рівняння, відповідно до теорії оптимальної лінійної фільтрації. За умови малої апостеріорної похибки, що описується дисперсією σ_t^2 , у марковській теорії можна застосовувати асимптотичні методи гаусової апроксимації. При цьому апостеріорний розподіл приблизно вважається нормальним:

$$w_t(\theta) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left\{-\frac{(\theta - m_t)^2}{2\sigma_t^2}\right\} \quad (3.77)$$

Вочевидь, це окремий випадок апроксимації (3.74) при $n=2$; $\varphi_1(\theta) = \theta$; $\varphi_2(\theta) = \theta^2$. Якщо підставити (3.77) у (3.70) або (3.71) і застосувати процедуру, аналогічну до тієї, яка була застосована до багатомірних розподілів, можна отримати рекурентні співвідношення, які показують як змінюється значення m_t апостеріорного середнього (його можна прийняти

за поточну оцінку $\theta^*(\theta)$ та дисперсії σ_t^2 вихідної похибки (результату фільтрації) із часом. У випадку одномірного дифузійного марковського процесу $\theta(t)$ і неперервного часу відповідні рівняння мають вигляд

$$\frac{dm_t}{dt} = A(m_t, t) + \sigma_t^2 b(t), \quad (3.78)$$

$$\frac{d\sigma_t^2}{dt} = B(m_t, t) - K(t)\sigma_t^4 + 2 \frac{\partial A(m_t, t)}{\partial m_t} \sigma_t^2, \quad (3.79)$$

$$b(t) = \frac{\partial F(y, \theta, t)}{\partial \theta}; \quad K(t) = -\frac{\partial^2 F(y, \theta, t)}{\partial \theta^2}.$$

Ці функції збігаються з тими, які входять у формул (3.63), (3.64) попереднього підрозділу. Функції ж A , B є коефіцієнтами рівняння Фоккера – Планка – Колмогорова

$$\frac{\partial w_t(\theta)}{\partial t} = -\frac{\partial}{\partial \theta} [A(\theta, t)w_t(\theta)] + \frac{1}{2} \frac{\partial^2}{\partial \theta^2} [B(\theta, t)w_t(\theta)] \quad (3.80)$$

яке є у цьому випадку конкретизацією рівняння (3.67). У окремому випадку, коли функція $[A(\theta, t)]$ є лінійною по відношенню до θ , а $B(\theta, t)$ не залежить від θ (при цьому марковський процес одночасно є гауссовим), рівняння (3.78), (3.79) описують метод пошуку оптимальної оцінки, яка точно співпадає з методом, розглянутим у підрозділі 3.5. У більш загальному випадку ланки згладжування оптимального вимірювача є нелінійними, що ілюструється схемою (рис. 3.4), яка відповідає рівнянням (3.78), (3.79).

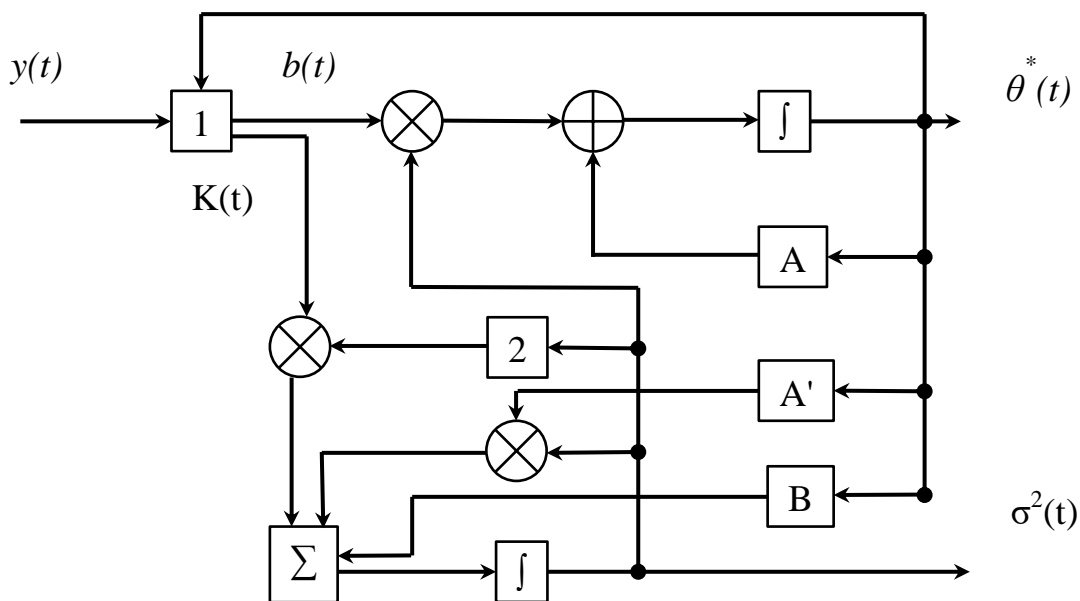
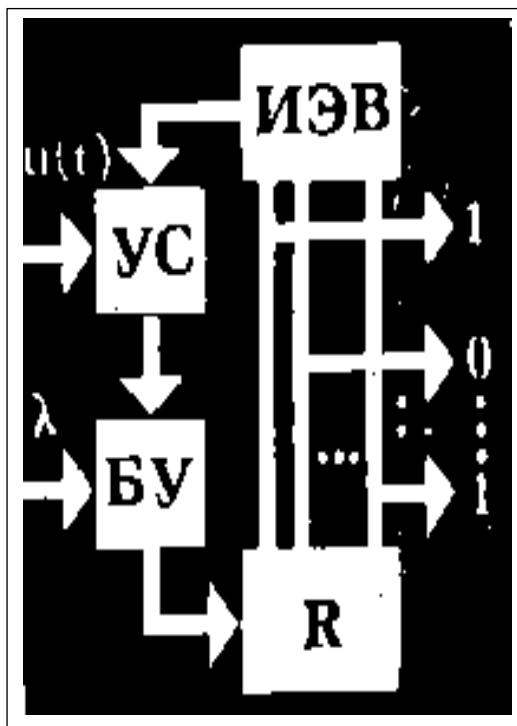


Рисунок 3.4 – Схема нелінійної фільтрації марковського процесу:
1 – оптимальний дискримінатор; 2 – блок зведення до квадрату; A, A', B – нелінійні безінерційні блоки з характеристиками $A(m_t, t)$, $A'(m_t, t)$, $B(m_t, t)$.



СУЧАСНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В РАДІОТЕХНІЧНИХ СИСТЕМАХ

- Базові поняття сучасної системотехніки
- Ймовірнісні моделі вимірюваних параметрів
- Оптимальне оцінювання параметрів
- Нелінійна оптимальна фільтрація

4 БАЗОВІ ПОНЯТТЯ СУЧАСНОЇ СИСТЕМОТЕХНІКИ

Існуючі на сьогоднішній день методи аналізу та синтезу радіотехнічних систем умовно можуть бути розбиті на три основні групи.

До *першої групи* відносять операторні методи, що базуються на перетвореннях Фур'є, Лапласа або z -перетвореннях, на основі яких синтез систем виконується широко відомим методом проб та помилок.

До *другої групи* відносять методи, у яких для оцінки якості системи використовується інтегральна квадратична або середньоквадратична похибка. Ці дві групи методів лежать у основі, так званої, класичної теорії радіотехнічних систем, основи якої уже було викладено у попередніх підрозділах даного навчального посібника.

Поява швидкодійних цифрових обчислювальних машин та пристроїв призвела до майже революційних змін у методах аналізу та синтезу радіотехнічних систем [15]; з'ясувалась непридатність старих методів та і самого підходу до синтезу тим самим був заданий потужний поштовх до розробки нового погляду на проблему синтезу та методів її розв'язування. Унаслідок цього, для побудови оптимальної структури синтезованої системи потрібно було виконувати трудомісткі обчислення за певними алгоритмами або програмами з великою кількістю обмежень. Практична реалізація таких обчислень обов'язково передбачала всебічне застосування цифрових обчислювальних засобів. Ця – *третья група* методів, для характеристики систем використовує базові поняття сучасної системотехніки, зокрема поняття системи у сучасному тлумаченні, поняття стану, простору станів системи та змінної стану. Основна ідея цієї групи методів полягає у тому, що вибирається або створюється функціонал від функції змінних стану та входних діянь, накладаються, з фізичних міркувань, обмеження на ці функції і за допомогою одного з методів варіаційного обчислення (динамічного програмування, принципу максимуму Понтрягіна або множників Лагранжа) визначається оптимальна стратегія, яка максимізує або мінімізує введений функціонал. Останніми роками цей підхід отримав загальне визнання і став основним методом синтезу сучасних радіотехнічних систем [19].

4.1 Основні засади загальної теорії систем та системного підходу

Поняття системного підходу. Системно-теоретичний підхід завжди являв собою невід'ємну частину методів, які використовувались у інженерній практиці. Це було пов'язано, головним чином з тим, що на відміну від інших наукових дисциплін, суть роботи інженера полягає не в тому щоб відкривати нові функціональні зв'язки фізичної природи, а в тому щоб розробляти системи, що складаються з сукупності багатьох компонентів та заставити працювати їх як єдине ціле. Системно-теоретична точка зору пов'язана з макроскопічним описом природних явищ і використовує, як

правило, тільки ті характеристики, які необхідні для побудови фізично коректної та працездатної системи. Прикладом системно-теоретичного мислення може слугувати діяльність інженера-електрика, який застосовує закони Ома, а не закони електродинаміки для розрахунків електричних кіл зі зосередженими параметрами.

Науково-технічна революція 1960–1970 років та поява цифрових обчислювальних машин застали інженера в такому стані творчої діяльності, коли потреба у розвитку та формалізації системно теоретичного підходу стала нагальною. Ті системи, які у подальшому були названі «великими системами» охоплювали усі аспекти інженерної науки і потребували розвитку певної абстрактної теорії, яку можна було б застосовувати до кожної інженерної дисципліни, не вдаючись у суть фізичних процесів. Математика дала формальний апарат для цієї нової наукової дисципліни, яка була названа загальною теорією систем або системотехнікою.

Для того щоб зрозуміти потенційні можливості загальної теорії систем та визначити сферу можливих застосувань доцільно дати означення базових понять, які використовує фахівець у галузі системотехніки. На відміну від інших випадків, такий підхід вкрай необхідний, оскільки ці означення повинні бути функціональними і призначаються для вирішення нагальних проблем. Такі поняття досліджувались багатьма вченими, зокрема, основоположником загальної теорії систем Л. Берталанфі. Біля витоків теорії систем також були американські вчені Р. Калман, М. Месарович та вітчизняні — А. М. Колмогоров, В. М. Глушков, М. М. Мойсєєв. Загальна теорія систем сформувалася як науковий напрям на початку другої половини ХХ ст. Вагомий внесок у її становлення зробили К. Боулінг, Р. Жерар, А. Хол, В. Афанасьєв, І. Блауберг, П. Анохін, В. Садовський. Створене в 1954 р. у США «Товариство досліджень в області загальної теорії систем» започаткувало видання щорічника «General Systems». У 1959 р. був створений центр системних досліджень при Цейсівському технологічному інституті, а в 1963 р. — Інститут системних досліджень.

Система як наукова категорія. Один з найкращих способів пояснення основних ідей загальної теорії систем полягає в описі її базових та фундаментальних концепцій, наприклад «система», «стан системи», «змінна стану». Поняття «система» запроваджене Аристотелем (у перекладі з грецької означає ціле, складене з частин, об'єднання), який відзначив, що сума частин характеризується тим, що становище цих частин в системі не утворює відмінностей. Але там, де виникають такі відмінності, має місце ціле. Заради справедливості, потрібно зауважити, що поняття система використовувалось і до Аристотеля. Спектр значень слова «система» у грецькій мові досить широкий: поєднання, організм, будова, організація, спілка, устрій, керівний орган. Багатозначність поняття «система» дозволяє визначити цим словом широке коло різноякісних явищ, які, проте, мають дещо спільне. Увібравши у себе сутність таких важливих понять, як «порядок», «організація», «цілісність», вона (система) не зводиться в повній

мірі ні до одного з них і, за висловами багатьох вчених поняття «система» стало певною мірою аксіоматичним. Таке поняття системи, як і будь-яке інше наукове поняття, є результатом логічного абстрагування і виступає значною мірою як ідеальний об'єкт, який формується у свідомості (уяві) дослідника. У визначенні категорії «система» досить часто спостерігається намагання надати йому чіткість та конкретність, прив'язавши до певної галузі досліджень. Наприклад, французький філософ Кондільяк говорив про систему, як про певний порядок розташування різних частин мистецтва або науки. Відомий філософ І. Кант стверджував, що система — це єдність різноманітних знань, які об'єднані однією ідеєю. Гегель підкреслював, що реальність має системний характер. Застосування системного світогляду сприяло видатним науковим відкриттям (Ч. Дарвін, Д. Менделєєв, В. Вернадський). У нинішній час системний стиль мислення став панівним. Аналіз визначення системи містять багато наукових праць з найрізноманітніших галузей, від філософії до ядерної фізики, від мистецтвознавства до технології виробництва, проте для початку доцільно підкреслити необхідність чіткого розмежування гносеологічного та онтологічного підходів до визначення системи. Гносеологічний (пізнавальний) підхід є відображенням у свідомості людини (суб'єктивним образом) реального матеріального утворення (системи) з метою його пізнання. Онтологічний підхід відображає реально існуючий взаємозв'язок об'єктів матеріального світу. Існує величезна кількість термінологічних визначень категорії «система». Наведемо найпоширеніші з них:

- система — цілісна взаємозалежна безліч об'єктів;
- система — цілісна безліч об'єктів (елементів), пов'язаних між собою взаємними відносинами;
- система — порядок (план, класифікація), згідно з яким розташовується група понять для утворення єдиного цілого;
- система — сукупність взаємозалежних, певним чином організованих і взаємодіючих елементів;
- система — організована безліч структурних елементів, що взаємопов'язані і виконують певні функції;
- система — сукупність взаємозалежних елементів, відособлена від середовища і взаємодіюча з ним як ціле.

Однак ці означення не є функціональними, так як не містять достатньої і конкретної математичної форми, необхідної для створення математичної моделі. Більш точно з математичної точки зору визначення терміну «система» дано у роботі Л.А. Заде [19].

Визначення 1. Система характеризується набором впорядкованих пар функцій часу, які являють собою її входи та виходи і визначають таким чином об'єкти абстрактного типу. Ці пари повинні задовольняти такій умові: якщо будь-яка пара з цих пар належить системі, то системі

належить також пара отримана шляхом звуження першо-початкової пари на меншу область визначення.

Таке визначення характеризує будь-який абстрактний об'єкт, а також будь-яку систему їх співвідношенням «вхід-вихід» і не залежить від конкретних фізичних властивостей притаманних відповідному процесові. Даним способом стає можливим досліджувати різноманітні системи у цілому виключно з точки зору співвідношень причинності між входом та виходом, що дозволяє одноманітно розглядати велику кількість різних проблем. Для того щоб описати структуру системи у математичних термінах, необхідно ввести ще два додаткових визначення.

Визначення 2. Поточний стан системи – це та мінімальна інформація про її минуле, яка необхідна для повного опису поведінки системи у майбутньому (тобто виходів систем) якщо поведінка її входів відома, починаючи з поточного моменту часу.

Визначення 3. Змінна стану – це та залежна від часу математична змінна, яка повністю визначає стан системи у будь-який момент часу.

Опираючись на ці аксіоми, систему можна описувати за допомогою інтегро-диференціальних рівнянь, які задають поведінку її станів або виходів у майбутні моменти часу, якщо задані або відомі поточний стан системи, минулі, поточні і майбутні входи системи. Такі системи називають *детермінованими*, а їх графічну та математичну модель позначають як на рис.4.1.

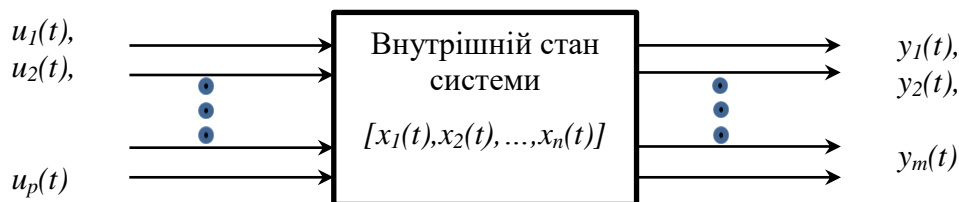


Рисунок 4.1– Опис системи в термінах змінних стану

Вхідні канали на цьому рисунку являють собою сукупність вхідних змінних або вхідних діянь $u_i(t)$, вихідні канали – сукупність вихідних змінних або реакцій системи $y_i(t)$, що дозволяють описати деякі аспекти поведінки системи, саме вони цікавлять спостерігача (дослідника) системи. Проміжні змінні або координати стану системи $x_i(t)$ відносяться до вмісту «чорного ящика» і таким чином вони сховані і недоступні для безпосереднього спостереження. Для зручності оперування з багатовимірними величинами сукупність вхідних змінних подамо у вигляді вектора входу розмірності p

$$\mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_p(t) \end{bmatrix}; \quad (4.1)$$

сукупність вихідних змінних – у вигляді вектора виходу розмірності m

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_m(t) \end{bmatrix}; \quad (4.2)$$

сукупність змінних стану – у вигляді вектора стану розмірності n

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}. \quad (4.3)$$

Згідно поняття векторного простору, множина усіх значень, які може прийняти вектор входу $\mathbf{u}(t)$ у момент часу t , утворює *простір вхідних сигналів системи*. Аналогічно, множина усіх значень, які може прийняти вектор виходу $\mathbf{y}(t)$ у момент часу t , утворює *простір вихідних сигналів системи*, і множина усіх значень, які може прийняти вектор стану системи $\mathbf{x}(t)$ у момент часу t , утворює *простір можливих станів системи*.

У будь-який момент часу t стан системи являє собою функцію від початкового стану системи $\mathbf{x}(t_0)$ та вектора вхідного сигналу $\mathbf{u}(t_0, t)$, тобто

$$\mathbf{x}(t) = \mathbf{F}[\mathbf{x}(t_0), \mathbf{u}(t_0, t)], \quad (4.4)$$

де \mathbf{F} – однозначна вектор-функція своїх аргументів.

Вектор виходу у будь-який момент часу t також являє собою деяку іншу вектор-функцію від функції входу $\mathbf{u}(t_0, t)$ та вектору початкового стану системи $\mathbf{x}(t_0)$ і може бути записаний у формі

$$\mathbf{y}(t) = \mathbf{G}[\mathbf{x}(t_0), \mathbf{u}(t_0, t)]. \quad (4.5)$$

Рівняння (4.4–4.5) часто називають рівняннями стану системи. У тому разі, коли система може бути описана звичайними диференціальними рівняннями, то (4.4–4.5) записують у загальній неявній формі

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{x}'(t) = \mathbf{F}[\mathbf{x}(t); \mathbf{u}(t)], \quad (4.6)$$

$$\mathbf{y}(t) = \mathbf{G}[\mathbf{x}(t); \mathbf{u}(t)]. \quad (4.7)$$

Для систем, які являють собою кінцеві автомати, рівняння стану приймають вигляд

$$\mathbf{x}(n) = \mathbf{F}[\mathbf{x}(n-1); \mathbf{u}(n-1)], \quad (4.8)$$

$$\mathbf{y}(n) = \mathbf{G}[\mathbf{x}(n); \mathbf{u}(n)]. \quad (4.9)$$

Якщо система може бути описана лінійними диференціальними рівняннями, то рівняння поточного стану системи зводяться до виразу

$$\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t); \quad (4.10)$$

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t), \quad (4.11)$$

де $A(t)$ – матриця коефіцієнтів системи; $B(t)$ – матриця комбінацій вхідних сигналів системи; $C(t)$ – матриця виходів системи; $D(t)$ – матриця обходу системи. Графічна модель системи, яка відповідає системі рівнянь (4.10–4.11) зображена на рис. 4.2. Вивід рівнянь стану системи, які повністю її характеризують, є першо-початковим етапом аналізу та синтезу сучасних інформаційно-вимірювальних систем.

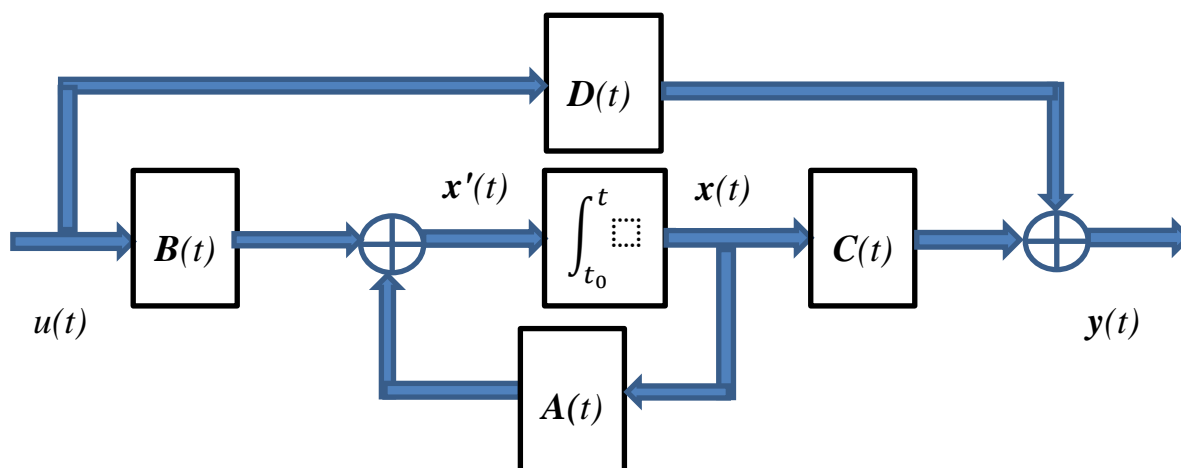


Рисунок 4. 2 – Подання математичної моделі системи в термінах змінних стану

У тому разі, коли можливі стани системи являють собою множину значень, заданих лише ймовірнісним розподілом, а входи є випадковими, то поведінку такої системи у майбутньому прийнято описувати лише з певною ймовірністю, і її називають *стохастичною*.

Якщо опис системи містить невідомі параметри, а вхідні сигнали є стохастичними і якщо система з плином часу зменшує ступінь своєї невизначеності, то таку систему називають *навчальною*.

У тому разі, коли система проявляє такі інтелектуальні здібності, як можливість прийняття складних рішень, планування та тощо, то її називають системою з *елементами штучного інтелекту*.

Серед технічних систем різноманітного цільового призначення особливе місце займають *системи автоматичного керування*, які завжди являють собою невід'ємну частину загальної теорії систем. Приклади використання таких систем можна виявити навіть на ранніх етапах людської цивілізації. Сучасний етап розвитку систем автоматичного керування датується серединою ХХ століття і продовжується нині завдяки піонерським роботам таких видатних вчених і інженерів як Вінер, Боде, Никольс, Найквіст, Калман, Ляпунов, Понтрягін, Попов, Фельдбаум та інші. Їх роботи склали єдину основу науки про керування, яка була втілена у найбільш визначних галузях сучасної техніки – від висадки людини на Місяць, вивчення космічного простору до керування біомедичними та біотехнічними пристроями з метою надання допомоги хворим та інвалідам. Задачі автоматичного ке-

рування розв'язувались у різний час за допомогою самих різноманітних математичних методів, які при цьому отримували різні назви у залежності від певних міркувань їх авторів. Для того щоб привести дані методи до єдиної точки зору, Підкомітет стандартів Наукового співтовариства із систем керування Інституту електротехніки і електроніки (ІЕЕЕ) вважає за доцільним дати наступні визначення.

Визначення 4. *Керування процесом являє собою таке втручання у перебіг процесу, яке забезпечує ефективне досягнення наміченої мети.*

Слова «ефективне досягнення» і «мета» тут використовуються у широкому розумінні і можуть означати, що виконуються деякі спеціальні умови, або можуть відповідати умовам оптимальності критерію якості у випадку оптимального керування. Перша спроба вирішити задачу керування у відповідності до вищенаведеного визначення полягала у використанні наперед розрахованої керуючої дії, яка дозволяла досягнути поставленої мети. Такий спосіб керування назвали *розімкненим або програмним керуванням*. Наступний крок, який став революційним у галузі керування, був зроблений у процесі конструювання електронних підсилювачів. При цьому *фактичний* вихід системи порівнювався з *бажаним (заданим)* виходом, а їх різниця використовувалась для керування системою. Такий метод керування отримав назву керування зі зворотним зв'язком і характеризується корекцією керуючої дії у темпі реального часу, що дозволило досягти наміченої мети без зовнішнього контролю за перебігом усього процесу. Такий підхід знаменував початок ери автоматизації і став визначальним для усіх сучасних технічних досягнень.

У подальшому, введення поняття *критерію якості* (у вигляді функціоналу), який містив у собі усі вимоги до системи, дозволило поставити і розв'язати задачу *оптимального керування*. Саме цей підхід отримав широке визнання завдяки своїй математичній елегантності та можливості застосування електронних обчислювальних машин для реалізації отриманих алгоритмів керування. Проте розвиток техніки, обумовлений розв'язком задач оптимального керування, поставив на порядок денний дослідження систем з більш складною структурою. Цей клас систем охарактеризувався необхідністю урахування у процесі керування непередбачуваного характеру вхідних діянь та можливістю випадкових змін параметрів системи з наперед відомими статистичними характеристиками. Це призвело до постановки задачі *стохастичного керування* (див. нижче наведене визначення), оптимальний розв'язок якої є одним з найпотужніших сучасних методів дослідження динамічних систем.

Визначення 5. *Задача стохастичного керування характеризується тим, що ступінь невизначеності, притаманний досліджуваному процесові, не може бути зменшений у ході перебігу процесу за рахунок статистичної обробки результатів вимірювань вихідних сигналів.*

На сьогоднішній день уже утворився достатньо великий клас систем, якість роботи яких погіршується через зміну зовнішніх умов за межами

діапазону, де можлива їх компенсація за допомогою регулятора в колі зворотного зв'язку. Інші системи подібного типу неадекватно працюють через занадто ідеалізоване або неточне моделювання їх поведінки. У обох випадках труднощі, що з'явилися можна пов'язати з наявністю невизначеності у описі системи або з недостатнім знанням сигналів, що діють на систему, статистичний опис яких наперед невідомий. Багато з таких систем можуть автоматично компенсувати відхилення своїх характеристик від номінальних за рахунок регулювання структури або параметрів, аналогічно тому як конструктор рекомендує видозміну системи, при отриманні додаткової інформації про її характеристики або поведінку. Цей новітній клас систем отримав назву *систем з самоорганізацією*. Проте методи їх досліджень виходять за межі даного навчального посібника.

4.2 Методи загальної теорії систем

Більшість актуальних задач, які являють практичний інтерес, пов'язано з системами, число станів яких велике. Труднощі такого типу зустрічаються в енергетичних системах, традиційних системах зв'язку та телекомунікацій, а також в соціально-економічних, екологічних системах, системах керування міським господарством, охорони здоров'я, які уже не відносяться до сфери технічних наук. Методи системотехніки, мабуть, зможуть допомогти у розв'язку вищезгаданих проблем, якщо врахувати успіхи застосування системного підходу у великих та складних програмах, наприклад, авіакосмічних. При системно-теоретичному підході для кожної системи підбирають підходящу математичну модель у розумінні визначення 1, яка не залежить від фізичної сутності системи. Аналіз математичної моделі системи дозволяє передбачувати її поведінку. На основі цих передбачуваних властивостей, у свою чергу, стає можливим вибирати відповідне керуюче діяння та випробовувати його на моделі. Потім, знайдені закони керування на математичній моделі доцільно переносити на реальний об'єкт. Формально варіант цієї процедури може бути представлений такою схемою.

1. Моделювання процесу.
2. Математичний аналіз.
3. Математичний синтез.
4. Випробовування керуючих дій.
5. Проектування реальної системи.

При цьому, слід зауважити, що тільки перший та останній пункти наведеної програми залежать від конкретних фізичних властивостей досліджуваного процесу. Останні три пункти можуть бути реалізовані за допомогою математичних методів та електронно-обчислювальної машини, що надає велику гнучкість запропонованому методу. Розглянемо більш детально опис усіх пунктів схеми.

Моделювання являє собою процес побудови математичної моделі досліджуваної фізичної системи. На етапі моделювання необхідна тісна

співпраця між експертом у даній конкретній галузі та фахівцем з теорії систем. Особливої уваги при моделюванні потребують системи, які містять невизначеності. У цьому випадку виконуються два етапи моделювання:

1. Структурна ідентифікація, яка містить у собі інтерпретацію відповідних фізичних законів.

2. Ідентифікація параметрів, суть якої полягає у визначенні значень функціональних параметрів системи.

Математичний аналіз являє собою процес вивчення зв'язку минулої, поточної та майбутньої поведінки системи із заданою структурою незалежно від її фізичної природи. Наприклад, для детермінованих систем можна обмежуватись аналізом відгуків на тестові вхідні сигнали. У випадку стохастичних систем доцільно скористатись методом оцінки станів. Для систем з високим рівнем невизначеності можна рекомендувати методи з самонавчанням. Оскільки процеси, що представляють практичний інтерес, як правило, характеризуються великою кількістю змінних станів, то для дослідження таких систем на сьогоднішній день розроблені різноманітні аналітичні методи, наприклад: ієрархічні методи; методи декомпозиції; методи мереж та мережевих графіків; теоретичні методи навчання.

Математичний синтез пов'язаний з вибором таких законів керування системою, які зможуть забезпечити досягнення бажаних робочих характеристик системи, що і є у багатьох випадках основною метою інженерного дослідження. Попередньо проаналізовану математичну модель примушують працювати бажаним чином за допомогою зовнішнього регулятора, який можна вмикати у систему розімкненого керування або у коло зворотного зв'язку. Нижче наводяться найпоширеніші методи визначення керуючої дії для динамічних систем: оптимізаційні методи математичного програмування; методи дослідження операцій; дослідження стійкості; методи стохастичного та дуального керування; методи керування з елементами самоорганізації, самонавчанням та штучного інтелекту.

Випробовування керуючих дій являє собою надзвичайно важливий етап системно-теоретичного підходу. Тут зазвичай перевіряють придатність розглянутих раніше процедур аналізу та синтезу. Потужні обчислювальні машини та комплекси суттєво спрощують етап апробації результатів синтезу. Найбільш поширені методи апробації передбачають: випробовування у зменшеному масштабі; моделювання на аналоговій або цифровій обчислювальній машині; проведення напівнатурного експерименту.

Проектування реальної системи пов'язано з використанням отриманої на математичній моделі інформації у першо-початковій реальній задачі, що вимагає, як зазначалося вище, обов'язкової співпраці між експертом у даній галузі та фахівцем з теорії систем. Цей етап системно-теоретичного підходу є завершальним.

5 ЙМОВІРНІСНІ МОДЕЛІ ВИМІРЮВАНИХ ПАРАМЕТРІВ

5.1 Вступ

Одним з найбільш важливих розділів дисципліни «Радіотехнічні системи» є розділ, присвячений виміру параметрів руху об'єктів. Як правило об'єкти, що розглядаються являють собою нелінійні динамічні системи, а похибки вимірювань та прикладені до об'єкту збурювання вважаються випадковими процесами. Тому задача оптимального проектування радіотехнічної системи за заданим критерієм якості є по суті задачею стохастичної теорії оптимального керування, яка ґрунтується на математичному апараті теорії статистичних рішень. Незважаючи на важливі результати, досягнуті за допомогою цієї теорії, практична реалізація одержаних алгоритмів керування нелійними системами наштовхується на суттєві труднощі, навіть при застосуванні сучасних мікро-ЕОМ [15]. Якщо ж об'єкт є лінійним і використовується квадратичний критерій якості, то можна отримати ефективні алгоритми розв'язку багатьох задач керування рухомими об'єктами. При цьому можна обмежуватись врахуванням лише кореляційних зв'язків між спостережуваними і керованими координатами, і таку теорію часто називають кореляційною теорією оптимального керування та оцінювання.

Істотна перевага кореляційної теорії керування та оцінювання полягає у тому, що з її допомогою відносно просто вдається синтезувати та реалізувати безпосередньо алгоритм роботи керуючого пристрою. При квадратичному критерії якості та лінійному об'єкті керування задача кореляційної теорії оптимального керування на підставі теореми розділення розбивається на дві взаємозалежні задачі:

- побудови найкращих лінійних оцінок випадкових координат об'єкту – задачу оптимальної лінійної фільтрації;
- синтезу оптимального регулятора, який керує цими оцінками.

Такий поділ відповідає поділу апаратури системи радіокерування на вимірювальне обладнання, що визначає координати об'єкту, і систему автоматичного керування рухом об'єкту. Причому, ці задачі можна розв'язувати незалежно одну від одної, незважаючи на те, що вони пов'язані спільною метою [6,19].

Матеріал даного навчального посібника відноситься, головним чином, до першої названої задачі, і присвячений сучасним методам синтезу оптимальних вимірювальних підсистем у радіотехнічних системах. У відповідності до цього напрямку досліджень, у наступних розділах викладені методи опису ймовірнісних моделей вимірювальних параметрів, методи розв'язку задач лінійної, нелінійної фільтрації векторних випадкових сигналів та процесів.

5.2 Математичні моделі випадкових сигналів та процесів

Значна кількість випадкових процесів, що підлягають подальшій обробці, може бути представлена як результат розв'язків певних стохастичних диференціальних або різницевих рівнянь. Перевагою такого підходу є те, що він придатний для опису як стаціонарних, так і нестаціонарних, як лінійних, так і нелінійних випадкових процесів, сформованих нелінійними (у загальному випадку) радіотехнічними системами.

Диференціальні рівняння використовуються у тому випадку, якщо процес описується у неперервному часі, різницеві – у дискретному часі. Можна розглядати дві еквівалентні форми запису диференціальних рівнянь: у вигляді диференціального рівняння n -го порядку або у вигляді n -мірної системи диференціальних рівнянь першого порядку. Другий варіант запису системи стохастичних лінійних диференціальних рівнянь у векторній формі має вигляд

$$\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{G}(t)\boldsymbol{\xi}(t) \quad (5.1)$$

де $\mathbf{x}(t)$ - векторна функція (вектор стану), що описує випадковий процес,

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \dots \\ x_n(t) \end{bmatrix};$$

$[x_1(t), x_2(t), \dots, x_n(t)]^T$ – складові процесу; \mathbf{A} – матриця розміром $n \times n$, $\boldsymbol{\xi}(t)$ – вектор білого шуму

$$\boldsymbol{\xi}(t) = \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \\ \dots \\ \xi_r(t) \end{bmatrix};$$

$[\xi_1(t), \xi_2(t), \dots, \xi_r(t)]^T$ – некорельовані між собою «білі» шуми з однаковими спектральними густинами; $\mathbf{G}(t)$ – матриця розміром $n \times r$. Рівняння (1) доповнюється вектором початкового значення $\mathbf{x}(0)$.

Векторне рівняння (1) легко складається за структурною схемою об'єкту або вигляду звичайного диференціального рівняння. Розглянемо приклади складання стохастичних диференціальних рівнянь.

Приклад 5.1

Якщо на вхід RC - ланки надходить «білий» шум $\boldsymbol{\xi}(t)$, то для вихідної напруги $u(t)$ можна скласти диференціальне рівняння

$$\mathbf{u}'(t) = \frac{1}{RC} \mathbf{u}(t) + \frac{1}{RC} \boldsymbol{\xi}(t).$$

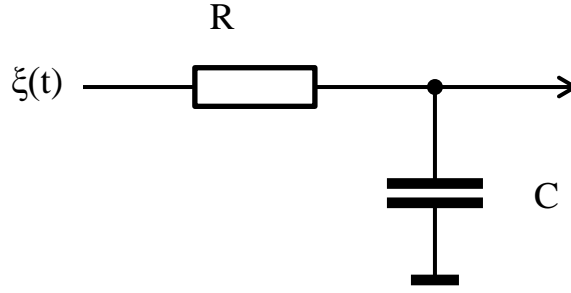


Рисунок 5.1– Структура RC ланки

Обчислимо автокореляційну функцію такого процесу. Розв'язок лінійного диференціального рівняння дається загально відомим виразом

$$u(t) = u(0)e^{-\frac{t}{RC}} + \frac{1}{RC} \int_0^t e^{-\frac{t-t_1}{RC}} \xi(t_1) dt_1,$$

При тривалому часі роботи можна вважати, що $t \rightarrow \infty$ і

$$u(t) \approx \frac{1}{RC} \int_0^t e^{-\frac{t-t_1}{RC}} \xi(t_1) dt_1.$$

Автокореляційна функція знаходиться шляхом осереднення за ансамблем реалізацій

$$\mathcal{M}[u(t)u(t')] = \left(\frac{1}{RC}\right)^2 \mathcal{M} \left\{ \int_0^t e^{-\frac{t-t_1}{RC}} \xi(t_1) dt_1 \int_0^{t'} e^{-\frac{t'-t_2}{RC}} \xi(t_2) dt_2 \right\}$$

Вважаючи припустимою операцію заміни місцями знаків осереднення та інтегрування, одержуємо

$$\mathcal{M}[u(t)u(t')] = \left(\frac{1}{RC}\right)^2 \int_0^t \int_0^{t'} e^{-\frac{t-t_1}{RC}} e^{-\frac{t'-t_2}{RC}} \mathcal{M}\{\xi(t_1) \xi(t_2)\} dt_1 dt_2.$$

Для білого шуму

$$\mathcal{M}\{\xi(t_1) \xi(t_2)\} = N\delta(t_1 - t_2).$$

Наявність дельта-функції у підінтегральному виразі дозволяє виконати інтегрування за однією змінною (наприклад t_2)

$$\begin{aligned} \mathcal{M}[u(t)u(t')] &= \left(\frac{1}{RC}\right)^2 \int_0^{\min(t,t')} e^{-\frac{t+t'-2t_1}{RC}} N dt_1 = \frac{N}{RC^2} e^{-\frac{t+t'}{RC}} \int_0^{\min(t,t')} e^{\frac{2t_1}{RC}} dt_1 = \\ &= \frac{N}{2RC} e^{-\frac{t+t'}{RC}} \left(e^{\frac{2\min(t,t')}{RC}} - 1 \right) = \frac{N}{2RC} \left(e^{\left| \frac{t-t'}{RC} \right|} - e^{-\frac{t+t'}{RC}} \right) \rightarrow_{t,t' \rightarrow \infty} \frac{N}{2RC} e^{-\frac{|t-t'|}{RC}}. \end{aligned}$$

Таким чином, дисперсія процесу $u(t)$ в усталеному режимі дорівнює $N/2RC$, а автокореляційна функція має вигляд експоненти. Такий процес прийнято називати експоненційно корельованим.

Приклад 5.2

Розглянемо диференційне рівняння, що описує рух центру мас твердого тіла під дією випадкової сили $f(t)$ типу «білий» шум. Прискорення маси твердого тіла, $a(t)$ визначається виразом

$$a = \frac{1}{m} f(t)$$

де m – маса. Враховуючи, що прискорення є першою похідною від швидкості $v(t)$ та другою похідною від пройденого шляху $s(t)$, можна скласти систему рівнянь

$$\begin{cases} s'(t) = v(t); \\ v'(t) = f(t)/m. \end{cases} \quad (5.2)$$

Систему рівнянь (5.2) можна записати у векторній формі:

$$\begin{bmatrix} s'(t) \\ v'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} f(t). \quad (5.3)$$

Якщо ввести вектор стану $\mathbf{x} = \begin{bmatrix} s(t) \\ v(t) \end{bmatrix}$ і позначити матриці

$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{G} = \begin{bmatrix} 0 \\ 1/m \end{bmatrix}$, тоді рівняння (3) можна записати у вигляді

$$\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{G}(t)f(t)$$

Відзначимо, що величини $s(t)$, $v(t)$ є інтегралами "білого" шуму і мають дисперсію, що наростає у часі. Таким чином, випадковий векторний процес у даному прикладі є нестационарним процесом, який складно досліджувати спектральними методами.

Приклад 5.3

Часто, у практичних задачах, прискорення об'єкту являє собою повільну флуктуаційну величину, яку важко і недоцільно описувати білим шумом. У подібних задачах прискорення можна зображати у вигляді експоненційно корельованого процесу з постійною часу τ . Модель з таким типом прискорення широко використовується для опису руху об'єктів, що маневрують, наприклад літаків.

Система диференціальних рівнянь, у цьому випадку, містить три рівняння: для дальності $d(t)$, швидкості $v(t)$ та прискорення $a(t)$

$$\begin{cases} d'(t) = v(t); \\ v'(t) = a(t); \\ a'(t) = (-1/\tau)a(t) + (1/\tau)\xi(t), \end{cases}$$

або у векторному записі

$$\begin{bmatrix} d'(t) \\ v'(t) \\ a'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1/\tau \end{bmatrix} \begin{bmatrix} d(t) \\ v(t) \\ a(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1/\tau \end{bmatrix} \xi(t). \quad (5.4)$$

Вираз (5.4) можна привести до вигляду (5.1), якщо ввести наступні позначення

$$\mathbf{x} = \begin{bmatrix} d(t) \\ v(t) \\ a(t) \end{bmatrix}; \quad \mathbf{d} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1/\tau \end{bmatrix}; \quad \mathbf{G} = \begin{bmatrix} 0 \\ 0 \\ 1/\tau \end{bmatrix}.$$

Рівняння (5.4) описує закон формування векторного випадкового процесу $x(t)$. Однак на практиці доступними для безпосереднього спостереження та вимірювання буває лише частина складових вектора $x(t)$. Тому необхідно описати процес вимірювань за допомогою рівняння спостережень. Для цього введемо m - мірний вектор спостережень $y(t)$, причому m дорівнює числу наявних вимірювальних пристроїв. Тоді $y(t)$ дорівнює

$$y(t) = H(t)x(t) + n(t), \quad (5.5)$$

де H – матриця спостережень розміром $m \times n$, $n(t)$ – вектор «білого» шуму спостережень, що описує похибки вимірювальних пристроїв.

Пояснимо складання рівняння спостережень на прикладі з літаком, що маневрує, вектор стану якого містить дальність, швидкість та прискорення і задається рівнянням (5.4). Вважаємо, що віддаль вимірюється радіодалькоміром, а швидкість доплерівським вимірювачем.

$$\begin{aligned} y &= \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} d(t) + n_1(t) \\ v(t) + n_2(t) \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} d(t) \\ v(t) \\ a(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix} = H(t)x(t) + n(t), \end{aligned}$$

де $y_1(t), y_2(t)$ – результати вимірювань віддалі і швидкості відповідно; $n_1(t), n_2(t)$ – похибки вимірювань віддалі і швидкості, відповідно. Матриця $H(t)$ може залежати від часу. Використання змінної матриці $H(t)$, дозволяє досліджувати вплив випадкових зникань корисного сигналу на точність вимірювань координат об'єкту у системі радіокерування. Наприклад, якщо у розглянутому вище випадку, через 100 секунд раптово припиняється вимір віддалі, то

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{при } t \leq 100 \text{ с,}$$

$$H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{при } t > 100 \text{ с.}$$

Розглянемо формулювання задачі фільтрації векторного випадкового процесу. Задані рівняння (5.1) і (5.5), які описують формування векторного випадкового процесу $x(t)$ і вектора вимірювань $y(t)$

$$x'(t) = A(t)x(t) + G(t)\xi(t); \quad (5.6)$$

$$y(t) = H(t)x(t) + v(t), \quad (5.7)$$

де $A(t), G(t), H(t)$ – матриці, що цілком певним чином залежать від часу; $\xi(t), v(t)$ – «білі» векторні шуми, що не залежать один від одного і мають нульові середні значення та наступні кореляційні матриці:

$$\mathcal{M}\{\xi(t_1) \xi(t_2)^T\} = I\delta(t_1 - t_2);$$

$$\mathcal{M}\{v(t_1) v(t_2)^T\} = R(t)\delta(t_1 - t_2);$$

де I – одинична матриця.

Передбачається, що $x(t_0)$ для початкового моменту часу є заданою величиною. Проте, вона може бути й випадковою. В такому разі передбачається, що вона має нормальний розподіл з середнім значенням $x_0 = x$ і кореляційною матрицею

$$P_0 = P.$$

Формування випадкових векторів $x(t)$ і $y(t)$ пояснюється на рис.5.2. Схемні зв'язки відмічені напівжирними лініями тому, що по них передаються векторні величини.

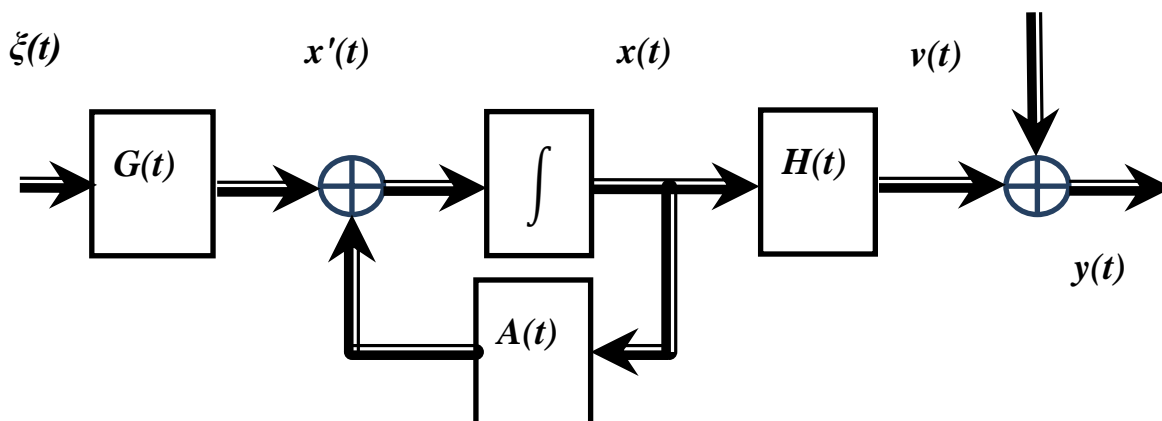


Рисунок 5.2 – Схема формування випадкових векторів $x(t)$ і $y(t)$

Задача оптимальної фільтрації полягає у тому, щоб на основі результатів спостережень $y(t)$ побудувати оцінку $x^*(t)$ з мінімальною середньоквадратичною похибкою.

Представлення величин x і y неперервному часу t доцільно у тому випадку, якщо обробка сигналів проводиться за допомогою аналогових пристроїв. При обробці інформації за допомогою цифрової ЕОМ введення інформації проводиться з певним тактовим інтервалом Δt , і безперервна величина $y(t)$ замінюється послідовністю відліків $y(1), y(2), \dots, y(k)$.

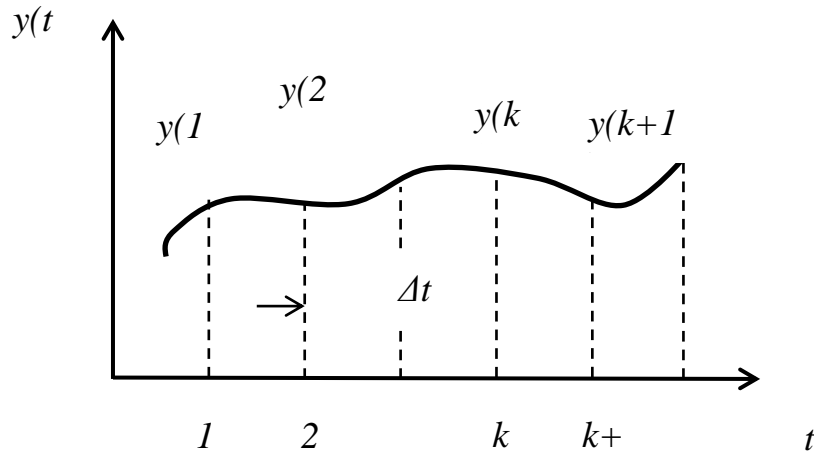


Рисунок 5.3 – Схема формування дискретних відліків

Рівняння руху реальних об'єктів звичайно задаються у вигляді системи диференціальних рівнянь, які визначають $x(t)$ і $y(t)$ у неперервному часі. Тому необхідний перехід від диференціального рівняння у неперервному часі до, так званого, різницевого рівняння у дискретному часі. Цей перехід необхідний ще й тому, що різницеві рівняння вирішуються чисельними методами набагато простіше, ніж диференційні. Складемо систему різницевого рівняння у такий спосіб. Задамо початкове значення $x(t_0)$ і розв'яжемо диференційне рівняння (5.1) для $t = t_1$, одержимо $x(1)$, потім візьмемо за початкове значення величину $x(1)$ і за допомогою цього ж диференціального рівняння знайдемо $x(2)$, і т.д.

Розв'язок лінійного диференціального рівняння (1) знаходиться як лінійна комбінація розв'язків однорідного та неоднорідного диференціальних рівнянь. Однорідне рівняння у цьому випадку має вигляд

$$x'(t) = A(t)x(t) \quad \text{при} \quad x(t_0) = x(0). \quad (5.8)$$

Розв'язок однорідного рівняння шукаємо у вигляді $x(t) = Ce^{-A(t) \cdot t}$, де C – довільна постійна величина, $C = x(0)$. Функцію $e^{-A(t) \cdot t}$ прийнято називати матричною експонентою, і вона може бути обчислена шляхом розв'язання експоненти у степеневий ряд. Якщо $A(t)$ не залежить від часу і $A(t) = A$, то можна знайти розв'язок однорідного рівняння за допомогою перетворення Лапласа. Перетворивши рівняння (5.8) за Лапласом, одержимо

$$px(p) - x(0) = Ax(p),$$

а потім знайдемо $x = [pI - A]^{-1}x(0)$. Якщо виконати обернене перетворення Лапласа, то отримаємо

$$x(t) = \Phi(t, t_0)x(0),$$

де $\Phi(t, t_0) = e^{-A(t-t_0)}$ – перехідна матриця станів, її отримують шляхом оберненого перетворення Лапласа величини $[pI - A]^{-1}$.

Розв'язок неоднорідного рівняння (5.1) виконується так само, як і розв'язок одномірного неоднорідного лінійного диференціального рівняння, але замість експоненти використовується перехідна матриця $\Phi(t, t_0)$:

$$x(t) = \Phi(t, t_0)x(0) + w(0), \quad (5.9)$$

де
$$w(0) = \int_{t_0}^t \Phi(t, \tau) G(\tau) \xi(\tau) d\tau.$$

Величина $w(0)$ є випадковою. Проаналізуємо її властивості. Слід підкреслити, що $w(0)$ має нормальний розподіл, оскільки вона отримана шляхом лінійного перетворення (інтегрування) випадкового процесу з нормальним розподілом ξ . Обчислимо моменти величини $w(0)$. Оскільки $\mathcal{M}\{\xi(t)\} = 0$, то і $\mathcal{M}\{w(0)\} = 0$.

$$\begin{aligned} \mathcal{M}\{w(0) w(0)^T\} &= \mathcal{M} \left\{ \int_{t_0}^t \Phi(t, \tau) G(\tau) \xi(\tau) d\tau \int_{t_0}^t \xi^T(\tau_1) G^T(\tau_1) \Phi^T(\tau_1, t) d\tau_1 \right\} \\ &= \int_{t_0}^t \int_{t_0}^t \Phi(t, \tau) G(\tau) \mathcal{M}\{\xi(\tau) \xi^T(\tau_1)\} G^T(\tau_1) \Phi^T(\tau_1, t) d\tau d\tau_1. \end{aligned} \quad (11.10)$$

Використовуючи властивість «білого» шуму $\mathcal{M}\{\xi(\tau) \xi(\tau_1)^T\} = I\delta(\tau - \tau_1)$ виконаємо у цьому виразі інтегрування за змінною τ :

$$\mathcal{M}\{w(0) w^T(0)\} = Q(t_0, t) = \int_{t_0}^t \Phi(t, \tau) G(\tau) G^T(\tau) \Phi^T(\tau, t) d\tau.$$

Для визначення величини $Q(t_0, t)$ необхідно задатися матрицями $\Phi(t, \tau)$ і $G(\tau)$. Скориставшись рівнянням (5.9), можна скласти ланцюжок рівнянь, що зв'язують значення $x(0), x(1), \dots, x(k)$.

$$\begin{aligned} x(1) &= \Phi(t_0, t_1) x(0) + w(0); \\ x(2) &= \Phi(t_1, t_2) x(1) + w(1); \\ &\dots\dots\dots \\ x(k+1) &= \Phi(t_k, t_{k+1}) x(k) + w(k), \end{aligned} \quad (5.11)$$

де $x(k) = x(t_k); \quad w(k) = \int_{t_k}^{t_{k+1}} \Phi(\tau, t_{k+1}) G(\tau) d\tau.$

Випадкові процеси, що задаються диференційними рівняннями типу (5.1) і різницеvim рівнянням (5.12), є марковськими випадковими процесами, так як значення процесу у будь-який момент часу визначається лише значенням процесу у попереднім моменті часу. Якщо шуми та початковий стан мають нормальний розподіл, то густина розподілу ймовірності вектора стану $\mathbf{x}(k)$ являє собою багатовимірний нормальний розподіл.

Приклад 5.4.

Задана система рівнянь, що описує рух твердого тіла з випадковим прискоренням

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{G}\xi(t),$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \mathbf{G} = \begin{bmatrix} 0 \\ \sqrt{N} \end{bmatrix},$$

де N – спектральна густина прискорення. Визначимо матрицю Φ для цієї системи. Складемо матрицю $[p\mathbf{I} - \mathbf{A}]$.

$$[p\mathbf{I} - \mathbf{A}] = \begin{bmatrix} p & -1 \\ 0 & p \end{bmatrix}.$$

Обчислимо $[p\mathbf{I} - \mathbf{A}]^{-1}$.

$$[p\mathbf{I} - \mathbf{A}]^{-1} = 1/p^2 \begin{bmatrix} p & 1 \\ 0 & p \end{bmatrix} = \begin{bmatrix} 1/p & 1/p^2 \\ 0 & 1/p \end{bmatrix}.$$

Виконуючи обернене перетворення Лапласа для інтервалу квантування часу Δt , одержимо

$$\Phi = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}.$$

Визначимо матрицю \mathbf{Q}

$$\begin{aligned} \mathbf{Q} &= \int_{t_k}^{t_{k+1}} \Phi(\tau, t_{k+1}) \mathbf{G}(\tau) \mathbf{G}^T(\tau) \Phi^T(\tau, t_{k+1}) d\tau = \\ &= \int_{t_k}^{t_{k+1}} \begin{bmatrix} 1 & t_{k+1} - \tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \sqrt{N} \end{bmatrix} \begin{bmatrix} 0 & \sqrt{N} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t_{k+1} - \tau & 1 \end{bmatrix} d\tau = \\ &= \int_{t_k}^{t_{k+1}} \begin{bmatrix} t_{k+1} - \tau & t_{k+1} - \tau \\ t_{k+1} - \tau & 1 \end{bmatrix} N d\tau = N \begin{bmatrix} \Delta t^3 & \Delta t^2 \\ \Delta t^2 & \Delta t \end{bmatrix}. \end{aligned}$$

Позначимо $\mathbf{x}(t) = \begin{bmatrix} d(t) \\ v(t) \end{bmatrix}$, тоді диференційне рівняння прийме вигляд

$$\begin{bmatrix} \dot{d}(t) \\ \dot{v}(t) \end{bmatrix} = \begin{bmatrix} v(t) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \sqrt{N}\xi(t) \end{bmatrix}.$$

Схема, яка відповідає цьому диференційному рівнянню показана на рис.5а.

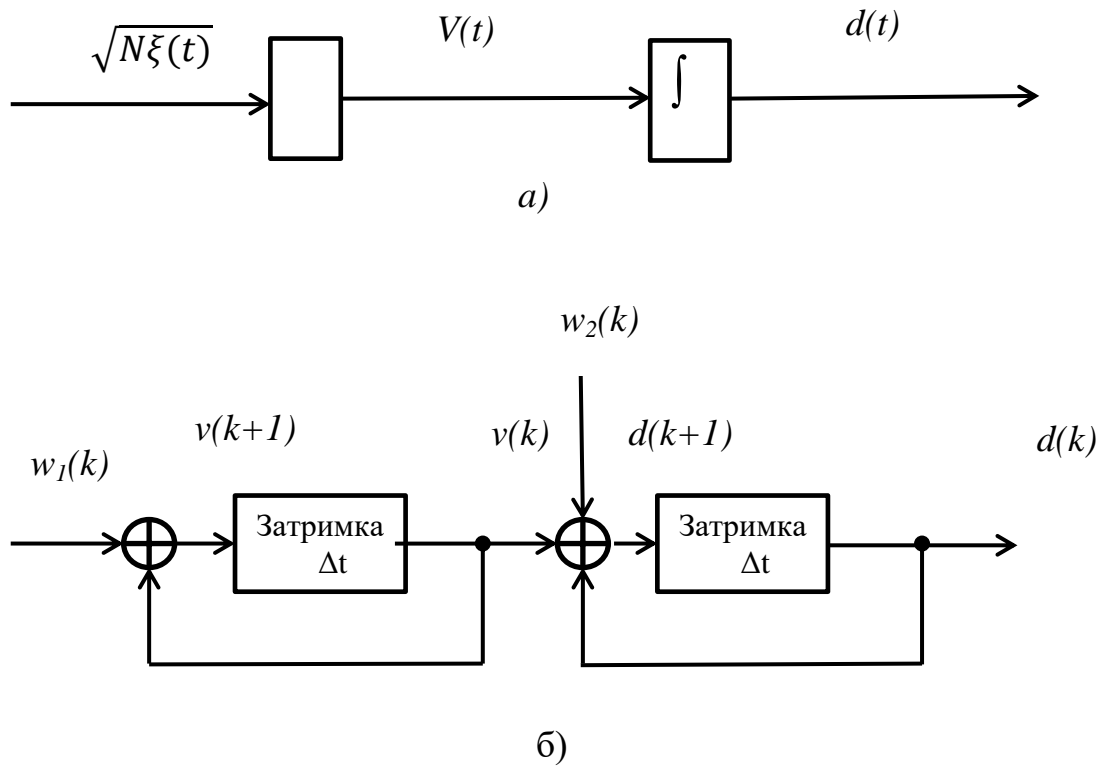


Рисунок 5.5 – Схема формування процесу з випадковим прискоренням

Тепер розглянемо систему різницевих рівнянь

$$\mathbf{x}(k + 1) = \Phi \mathbf{x}(k) + \mathbf{w}(k).$$

Враховуючи, що $\mathbf{x}(k + 1) = \begin{bmatrix} d(k + 1) \\ v(k + 1) \end{bmatrix}$, одержимо

$$\begin{bmatrix} d(k + 1) \\ v(k + 1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d(k) \\ v(k) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} = \begin{bmatrix} d(k) + \Delta t v(k) + w_1(k) \\ v(k) + w_2(k) \end{bmatrix}.$$

Системі різницевих рівнянь відповідає схема, яка зображена на рис. 5.5 б.

6 ОПТИМАЛЬНЕ ОЦІНЮВАННЯ ПАРАМЕТРІВ

Розглянемо розв'язок задачі фільтрації багатовимірного випадкового процесу, який описується у неперервному часі рівнянням (5.6). Процес вимірювання частини складових вектора $x(t)$ описується рівнянням (5.7). Задача фільтрації полягає в отриманні найкращої оцінки вектора $x^*(t)$ за результатами спостережень $y(t)$. За квадратичного критерію якості імпульсна реакція оптимального фільтра визначається відомим інтегральним рівнянням Вінера–Хопфа. Використовуючи це рівняння для векторного випадкового процесу, можна одержати диференціальне рівняння для найкращої оцінки:

$$\frac{dx^*(t)}{dt} = A(t)x^*(t) + K(t)[y(t) - H(t)x^*(t)]. \quad (6.1)$$

Реалізація алгоритму (6.1) може бути виконана у вигляді системи стеження.

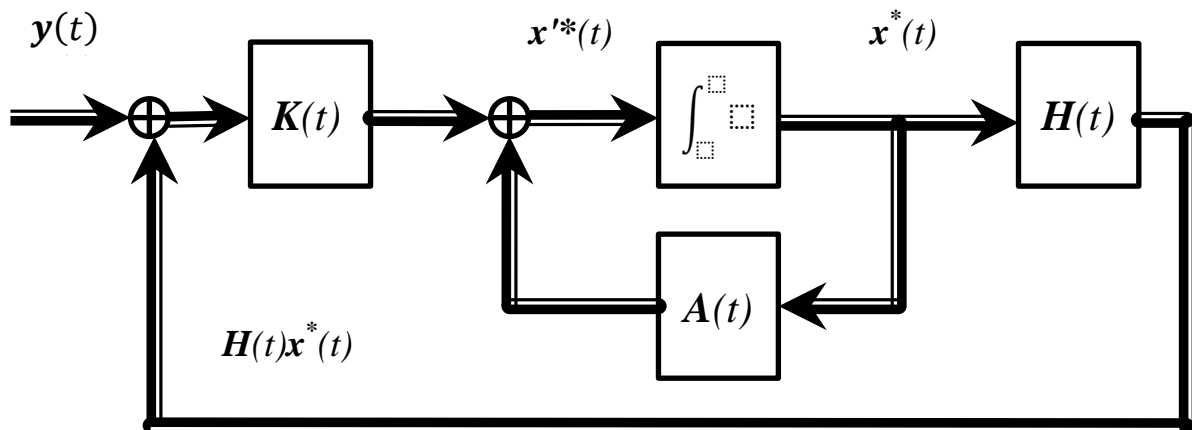


Рисунок 6.1 – Реалізація лінійного фільтра у неперервному часі

В оптимальній фільтрі виконується порівняння частини складових оцінки $H(t)x^*(t)$ з вихідним сигналом $y(t)$ і утворюється вектор сигналу похибки $[y(t) - Hx^*(t)]$. Ця величина потім подається до інтеграторів, які зберігають величину оцінки $x^*(t)$. Вага поправки визначається матричним коефіцієнтом підсилення $K(t)$. Величина матричного коефіцієнта підсилення $K(t)$ залежить від дисперсії похибок оптимальної фільтрації і визначається таким виразом:

$$K(t) = P(t)H^T(t)R^{-1}(t) \quad (6.2)$$

де $P(t) = E\{[x(t) - x^*(t)][x(t) - x^*(t)]^T\}$ є кореляційною матрицею похибок фільтрації, і яка визначається матричним диференціальним рівнянням Рікатті:

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) - \\ -\mathbf{P}(t)\mathbf{H}^T(t)\mathbf{R}^{-1}(t)\mathbf{H}(t)\mathbf{P}(t) + \mathbf{G}(t)\mathbf{G}^T(t) \quad (6.3)$$

з початковою умовою $\mathbf{P}(t_0) = \mathbf{P}(0)$. Диференційне матричне рівняння (6.3) еквівалентне системі з n^2 скалярних диференціальних рівнянь. Оскільки матриця $\mathbf{P}(t)$ – симетрична, то у системі залишається лише $(n^2 + n)/2$, різних рівнянь. При $n > 2$ розв'язок рівняння (6.3), як правило, виконується чисельними методами з використанням ЕОМ. У виразах (6.2) - (6.3) результати вимірювань $y(t)$ не використовуються, так що розрахунок оптимального коефіцієнта передачі може бути виконаний заздалегідь за апріорними даними. Алгоритм, що визначається виразами (6.1), (6.2) і (6.3), прийнято називати алгоритмом фільтра Калмана. Алгоритм фільтрації у вигляді диференціального рівняння (6.3) цікавий, головним чином, з теоретичної точки зору, тому що його апаратна реалізація надзвичайно складна. На практиці алгоритм фільтра Калмана реалізують за допомогою цифрової ЕОМ. Однак є певна незручність через те, що чисельний розв'язок проводиться у дискретному часі. Тому при обробці сигналів в ЕОМ бажано сформулювати задачу фільтрації для дискретного часу.

Розглянемо задачу оптимальної лінійної фільтрації для дискретного часу. Передбачається, що випадковий процес є марковським і задається векторним різницеvim рівнянням

$$\mathbf{x}(k+1) = \mathbf{\Phi}(k)\mathbf{x}(k) + \mathbf{w}(k), \quad (6.4)$$

де $\mathbf{\Phi}(k)$ - перехідна матриця, у загальному випадку залежна від k , $\mathbf{w}(k)$ – дискретний "білий" шум з нульовим середнім значенням і заданою кореляційною матрицею $\mathbf{Q}(k) = \mathbf{E}[\mathbf{w}(k)\mathbf{w}^T(k)]$, $\mathbf{x}(0)$ – наперед відома величина, або відомі її статистичні характеристики $\mathbf{E}[\mathbf{x}(0)]$ і $\mathbf{P}(0) = \mathbf{E}\{[\mathbf{x}(0) - \mathbf{x}^*(0)][\mathbf{x}(0) - \mathbf{x}^*(0)]^T\}$.

Вимірюванню підлягає лише частина складових вектора стану $\mathbf{x}(k)$ відповідно до рівняння

$$y(k) = \mathbf{H}(k)\mathbf{x}(k) + v(k) \quad (6.5)$$

де $\mathbf{H}(k)$ - матриця спостережень; $v(k)$ – дискретний "білий" шум з нульовим середнім значенням та заданою кореляційною матрицею $\mathbf{R}(k) = \mathbf{E}[v(k)v^T(k)]$. У теорії фільтрації випадкових процесів показано, що якщо початкове значення $\mathbf{x}(0)$ є детермінованою або випадковою величиною з нормальним розподілом, а шуми $\mathbf{w}(k)$ і $v(k)$ – гаусові, то найкраща оцінка вектора $\mathbf{x}^*(k)$ у сенсі мінімуму середнього квадрату похибки є лінійною. Однак, на практиці лінійну фільтрацію часто і успішно застосовують при розподілі початкового стану та завад, що відрізняються від нормального. В останньому випадку, фільтр залишається оптимальним лише в

класі лінійних фільтрів, а оптимальний фільтр, у загальному випадку, виявляється нелінійним. Для синтезу такого нелінійного фільтра, потрібні дані про розподіли, які не завжди доступні проектувальникові. При відсутності цих апріорних відомостей лінійний фільтр з мінімальною середньоквадратичною похибкою залишається єдиним можливим розв'язком.

Побудуємо рекурентну лінійну оцінку з мінімальною дисперсією. Припустимо, що відомі оцінка $\mathbf{x}^*(k)$ на k -му кроці та її кореляційна матриця похибок

$$\mathbf{P}(k) = E\{[\mathbf{x}(k) - \mathbf{x}^*(k)][\mathbf{x}(k) - \mathbf{x}^*(k)]^T\},$$

і потрібно побудувати оптимальну оцінку на $(k+1)$ -му кроці $\mathbf{x}^*(k+1)$. Доцільно розбити формування $(k+1)$ -ї оцінки на два етапи – етап екстраполяції на $(k+1)$ -й крок оцінки $\mathbf{x}^*(k)$ і етап фільтрації, на якому екстрапольована оцінка, корегується прийнятим сигналом $\mathbf{y}(k+1)$. Екстраполяція оцінки проводиться шляхом заміни значень $\mathbf{x}(k)$ і $\mathbf{w}(k)$ у виразі (6.4) їх оцінками $\mathbf{x}^*(k)$, $\mathbf{w}^*(k)$. Екстрапольована оцінка $\mathbf{x}^*(k+1/k)$ дорівнює

$$\mathbf{x}^*(k+1/k) = \Phi(k)\mathbf{x}^*(k/k) + \mathbf{w}^*(k).$$

Оскільки $\mathbf{w}(k)$ являє собою дискретний "білий" шум, то його значення не можна передбачати на основі попередніх спостережень, а тому отримана оцінка дорівнює її математичному сподіванню. Враховуючи, що $E[\mathbf{w}(k)] = 0$, одержимо

$$\mathbf{x}^*(k+1/k) = \Phi(k)\mathbf{x}^*(k/k). \quad (6.6)$$

Користуючись виразами (6.4) і (6.6), визначимо значення похибки екстраполяції $\mathbf{e}(k+1)$

$$\mathbf{e}(k+1) = [\mathbf{x}(k) - \mathbf{x}^*(k+1/k)] = \Phi(k)\mathbf{e}(k) + \mathbf{w}(k) \quad (6.7)$$

де $\mathbf{e}(k) = [\mathbf{x}(k) - \mathbf{x}^*(k/k)]$ – похибка фільтрації на k -ому кроці. Визначимо середнє значення цієї похибки

$$E[\mathbf{e}(k+1)] = \Phi(k)E[\mathbf{e}(k)].$$

Якщо оцінка $\mathbf{x}^*(k)$ була виконана так, що $E[\mathbf{e}(k)] = 0$ (зсув оцінки відсутній), то екстраполяція так само не має зсуву і $E[\mathbf{e}(k+1)] = 0$. Тепер обчислимо кореляційну матрицю похибок екстраполяції $\mathbf{P}(k+1/k)$.

$$\begin{aligned} \mathbf{P}(k+1/k) &= E[\mathbf{e}(k+1)\mathbf{e}^T(k+1)] = \\ &= E[\Phi(k)\mathbf{e}(k)\mathbf{e}^T(k)\Phi^T(k) + \mathbf{w}(k)\mathbf{w}^T(k)] = \Phi(k)\mathbf{P}(k/k)\Phi^T(k) + \mathbf{Q}(k). \end{aligned} \quad (6.8)$$

При обчисленні виразу (6.8) використовувалася умова відсутності кореляції між шумом $w(k)$ і похибкою $e(k)$, тому що $e(k)$ залежить тільки від значень $w(k-1)$, $w(k-2)$,

Тепер розглянемо формування оптимальної оцінки на етапі фільтрації. Оскільки операція фільтрації лінійна, то оцінка дорівнює лінійній комбінації величин $x^*(k+1/k)$ і $y(k+1)$

$$x^*(k+1/k+1) = K_1 x^*(k+1/k) + K_2 y(k+1), \quad (6.9)$$

де K_1 , K_2 постійні матричні коефіцієнти. Оптимальність оцінки забезпечується відповідним вибором коефіцієнтів. При виборі коефіцієнтів враховуються дві вимоги: умова відсутності зсуву оцінки та вимога мінімізації середнього квадрата похибки. Умова відсутності зсуву оцінки полягає в тому, щоб

$$E[x(k+1) - x(k+1/k+1)] = 0.$$

Користуючись рівняннями (6.4) та (6.9) можна отримати

$$\begin{aligned} E[x(k+1) - K_1 x^*(k+1/k) - K_2 y(k+1)] = \\ = E\{[I - K_2 H(k+1)]x(k+1) - K_1 x^*(k+1/k)\}. \end{aligned}$$

При відсутності зсуву цей вираз повинний бути нульовим, а це можливо лише тоді, коли виконується умова

$$[I - K_2 H(k+1)]E\{x(k+1)\} = K_1 E\{x^*(k+1/k)\}.$$

Оскільки $E\{x(k+1)\} = E\{x^*(k+1/k+1)\}$, то $K_1 = [I - K_2 H(k+1)]$.

Таким чином, вираз (6.9) приймає вигляд

$$\begin{aligned} x^*(k+1/k+1) &= [I - K_2 H(k+1)]x^*(k+1/k) + K_2 y(k+1) = \\ &= x^*(k+1/k) + K_2 [y(k+1) - H(k+1)x^*(k+1/k)]. \end{aligned} \quad (6.10)$$

Оцінка, що надається виразом (6.10), не має зсуву при будь-якому K_2 . Тепер визначимо оптимальне значення матричного коефіцієнта підсилення K . Для цього обчислимо кореляційну матрицю похибок фільтрації $P(k+1/k+1)$, користуючись формулою (6.10):

$$\begin{aligned} P(k+1/k+1) &= \\ &= E\{[x(k+1) - x^*(k+1/k+1)][x(k) - x^*(k+1/k+1)]^T\} = \\ &= [I - K_2 H(k+1)] P(k+1/k) [I - K_2 H(k+1)]^T + K_2 R(k+1) K_2^T \end{aligned} \quad (6.11)$$

Слід підкреслити, що вивід виразу (6.11) базувався на припущенні відсутності кореляції між шумом $v(k+1)$ та похибками екстраполяції $e(k+1)$. Оскільки похибка $e(k+1)$ являє собою векторну величину, то виникає питання про вибір скалярної величини, яка б характеризувала величину вектора $e(k+1)$, тобто норму вектора $e(k+1)$. Найпростішим типом такої норми є сума квадратів складових вектора $e(k+1)$, яка чисельно дорівнює сліду кореляційної матриці, який позначимо через $\text{Sp } P$.

При оптимізації коефіцієнта K_2 слід мінімізувати $\text{Sp } P$. Щоб знайти мінімум по K_2 , необхідно обчислити градієнт по K_2 від $\text{Sp } P$ і знайти таке значення K_2 , при якому, градієнт би дорівнював нулю.

Розглянемо обчислення градієнта сліду квадратичної форми $\text{Sp } KRK^T$.

$$\text{Grad } [\text{Sp } KRK^T] = e_{11} \frac{\partial}{\partial K_{11}} \text{Sp } KRK^T + e_{12} \frac{\partial}{\partial K_{12}} \text{Sp } KRK^T + \dots + e_{1n} \frac{\partial}{\partial K_{1n}} \text{Sp } KRK^T,$$

где e_{ij} – одиничні вектори, що відповідають елементу K_{ij} матричного коефіцієнту підсилення. Виконавши диференціювання одного з доданків, з урахуванням симетрії матриці R одержимо:

$$e_{ij} \frac{\partial}{\partial K_{ij}} \text{Sp } KRK^T = 2 \text{Sp} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & e_{ij} & \dots & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} RK^T$$

У такий спосіб обчислення градієнту квадратичної форми дає вираз

$$\text{Grad } [\text{Sp } KRK^T] = 2 \text{Sp } ERK^T,$$

де
$$E = \begin{bmatrix} e_{11} & \dots & e_{1n} \\ e_{21} & \dots & e_{2n} \\ e_{n1} & \dots & e_{nn} \end{bmatrix}$$

Скориставшись таким правилом обчислення градієнту, знайдемо умову, при якій матричний коефіцієнт підсилення $K(k+1)$ буде оптимальним:

$$2 \text{Sp } E \{-H(k+1)P(k+1/k) [I - K_2(k+1)H(k+1)]^T + R(k+1)K_2^T(k+1)\} = 0.$$

Оскільки матриця E заповнена одиничними векторами, то рівність нулю може вийти лише за умови

$$-H(k+1)P(k+1/k) [I - K_2(k+1)H(k+1)]^T + R(k+1)K_2^T(k+1) = 0$$

Після алгебраїчних викладок одержуємо

$$K_{2opt}(k+1) = K(k+1) = \\ = P(k+1/k)H^T(k+1)[H(k+1)P(k+1/k)H^T(k+1) + R(k+1)]^{-1}. \quad (6.12)$$

Відзначимо, що при $K = K(k+1)$ вираз (6.11) для кореляційної матриці P може бути спрощений

$$P(k+1/k+1) = [I - K(k+1)H(k+1)]P(k+1/k). \quad (6.13)$$

Таким чином, остаточний вигляд оптимальної лінійної оцінки буде описуватись виразом:

$$x^*(k+1/k+1) = \Phi(k)x^*(k/k) + K(k+1)z(k+1), \quad (6.14)$$

де $z(k+1) = y(k+1) - H(k+1)\Phi(k)x^*(k/k)$, $x^*(0/0) = E[x(0)]$.

Величина $z(k+1)$ являє собою ту нову, додаткову інформацією, яка міститься у повідомленні $y(k+1)$. Саме з цієї причини процес $z(k)$ називають "обновлювальним" процесом. Можна показати, що у оптимальному фільтрі величина $z(k)$ являє собою "білий" дискретний шум. Для обчислення оцінки $x^*(k/k)$ необхідно обчислити послідовність оптимальних коефіцієнтів підсилення $K(k)$. Для розрахунків цієї послідовності використовуються вирази (6.11), (6.12) і (6.13). Оскільки $K(k)$ не залежить від результатів спостережень $y(k)$, то він може бути розрахований заздалегідь за апіорним даними. Алгоритм фільтрації, що описується формулами (6.14), (6.11)–(6.13), прийнято називається дискретним фільтром Калмана. Цей алгоритм зручний для розрахунків на ЕОМ, тому що формули для обчислення $x^*(k/k)$ і $P(k/k)$ є рекурентними і у пам'яті ЕОМ не потрібно зберігати їх попередні значення. Фільтр Калмана можна зображати у вигляді системи стеження (рис. 6.2).

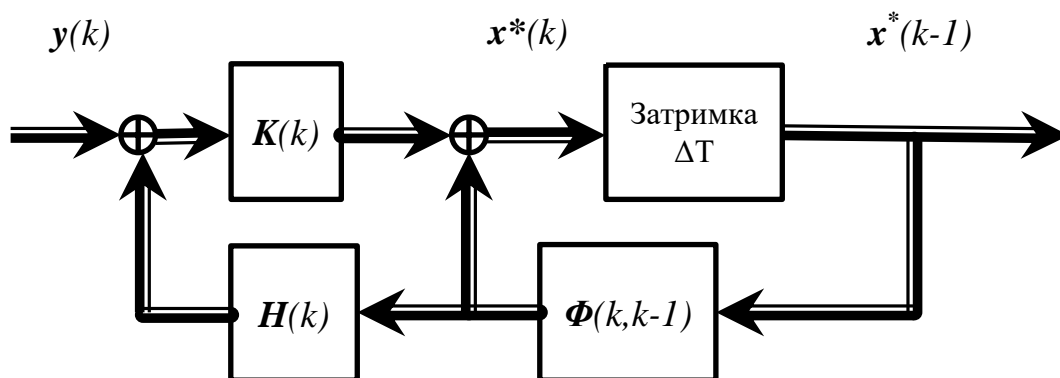


Рисунок 6.2 – Структура дискретного фільтра Калмана

Порівнюючи з моделлю формування сигналу можна відзначити, що структура фільтра визначається типом моделі випадкового процесу.

Приклад 6.1

Розглянемо приклад фільтрації випадкового процесу, заданого різницеvim рівнянням

$$\mathbf{x} = \begin{bmatrix} d(k+1) \\ v(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d(k) \\ v(k) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_{12}(k) \end{bmatrix},$$

причому вважаються відомими величини $\mathbf{E}[\mathbf{x}(0)]$, $\mathbf{P}(0/0)$, $\mathbf{Q}(k) = \mathbf{E}[\mathbf{w}(k)\mathbf{w}(k)]$.

Рівняння спостережень має вигляд

$$y(k) = \mathbf{H}\mathbf{x}(k) + n(k),$$

де $\mathbf{H} = [1 \ 0]$, і відомо, що $\mathbf{R}(k) = \mathbf{E}[n^2(k)]$.

Вектор стану $\mathbf{x}(k)$ можна розглядати як вектор, що описує рух об'єкту за віддалю $d(k)$ зі швидкістю $v(k)$. Рівняння спостереження дає значення відліків віддалі $y(k)$.

Побудуємо оптимальний фільтр відповідно до виразів (6.6) і (6.14). Оцінка, екстрапольована на $(k+1)$ -й інтервал часу, дорівнює

$$\mathbf{x}^*(k+1/k) = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d^*(k/k) \\ v^*(k/k) \end{bmatrix} = \begin{bmatrix} d^*(k/k) + \Delta t v^*(k/k) \\ v^*(k/k) \end{bmatrix}.$$

Таким чином, значення $d^*(k+1/k)$ отримується шляхом додавання поправки $\Delta t v^*(k/k)$ до попереднього значення віддалі $d^*(k/k)$. Зміни оцінки швидкості $v^*(k/k)$ при екстраполяції не відбувається, тому що в розглянутій задачі прискорення об'єкта є "білим" шумом і його екстрапольоване значення дорівнює нулю. Оцінка після обробки відліку віддалі $y(k+1)$ дорівнює

$$\mathbf{x}^*(k+1/k+1) = \begin{bmatrix} d^*(k+1/k) + K_1 z(k+1) \\ v^*(k+1/k) + K_2 z(k+1) \end{bmatrix},$$

де $z(k+1) = y(k+1) - \mathbf{H}\mathbf{x}^*(k+1/k) = y(k+1) - d^*(k+1/k)$.

Схема оптимального фільтра показана на рис. 6.3. Ця схема відрізняється від розглянутої вище схеми (рис. 6.2) тим, що в ній розкриті векторні зв'язки.

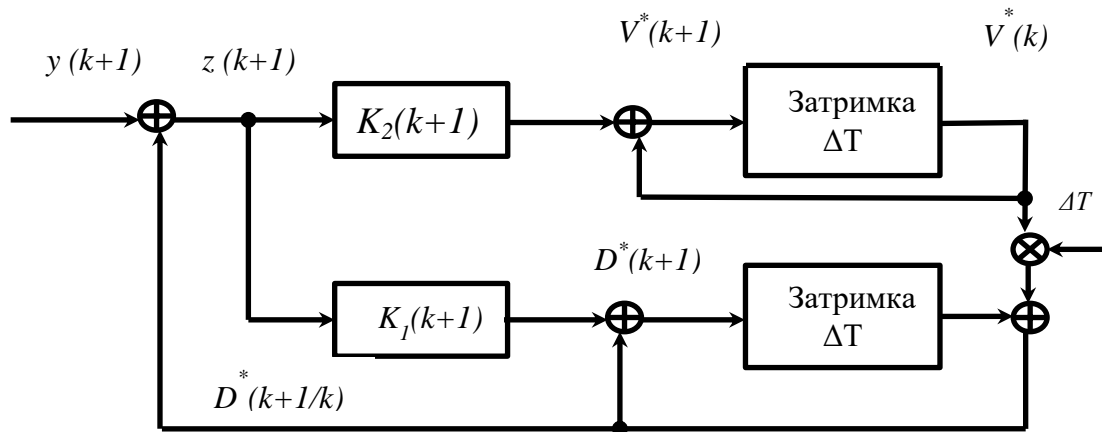


Рисунок 6.3– Структура дискретного фільтра для прикладу 6.1

Оптимальний фільтр можливо розглядати як систему стеження з двома дискретними інтеграторами, які утворюються замкненими контурами з елементами затримки на інтервал дискретизації Δt . Дискретні інтегратори використовуються для зберігання оцінок швидкості $v^*(k/k)$ і віддалі $d^*(k/k)$. Екстраполяція оцінки віддалі на $(k+1)$ -й такт проводиться шляхом підсумовування поправки $\Delta t v^*(k/k)$ до значення віддалі $d^*(k/k)$. Потім екстрапольовані значення віддалі порівнюються з результатами вимірювань віддалі $y(k+1)$, а їх різниця використовується для корегування значень $d^*(k/k)$ і $v^*(k/k)$. Вага поправок визначається коефіцієнтами підсилення $K_1(k+1)$ та $K_2(k+1)$, які залежать, у свою чергу, від кореляційної матриці похибок $\mathbf{P}(k/k)$. При великих похибках фільтрації коефіцієнти підсилення великі, і навпаки. Як правило, значення коефіцієнтів підсилен-

ня зменшуються із часом, тому що помилки в усталеному є значно меншими ніж на етапі введення в режим стеження. Покажемо як змінюються коефіцієнти підсилення на чисельному прикладі. Припустимо, що задані величини $d(0)$, $v(0)$, $P(0/0) = \begin{bmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}$.

Величина σ_d^2 є дисперсією початкової похибки вимірювань віддалі, яка визначається властивостями схеми пошуку сигналу по віддалі. Величина σ_v^2 при відсутності апріорних даних про швидкість визначається швидкістю руху об'єкта. Якщо швидкість руху має рівномірний розподіл у межах $[-v_{max}, +v_{max}]$, то $\sigma_v^2 = (2/3) v_{max}^2$.

Для спрощення розрахунків вважаємо, що $Q(k) = 0$, $R(k) = 0$. Зауважимо, що рівняння (6.4) за прийнятих припущень описує рух з постійною початковою швидкістю $v(0)$ з початкового стану $d(0)$. На першому такті з рівнянь (6.12) і (6.13) одержуємо

$$P = \begin{bmatrix} \sigma_d^2 + \sigma_v^2 \Delta t^2 & \sigma_v^2 \Delta t \\ \sigma_v^2 \Delta t & \sigma_v^2 \end{bmatrix}; K(1) = \begin{bmatrix} 1 \\ \frac{-\sigma_v^2 \Delta t^2}{\sigma_d^2 + \sigma_v^2 \Delta t^2} \end{bmatrix}; P = [I - K(1)HP(1/0)] = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\sigma_d^2 \sigma_v^2}{\sigma_d^2 + \sigma_v^2 \Delta t^2} \end{bmatrix}.$$

Відзначимо, що дисперсія помилки виміру віддалі дорівнює 0, оскільки виміри відбуваються без похибок. Продовжуючи обчислення далі, отримуємо

$$P(2/1) = \frac{\sigma_d^2 \sigma_v^2}{\sigma_d^2 + \sigma_v^2 \Delta t^2} \begin{bmatrix} \Delta t^2 & \Delta t \\ \Delta t & 1 \end{bmatrix}.$$

Похибки екстраполяції на другому такті відмінні від нуля, тому що ще невідомо точне значення швидкості. Значення $K(2) = \begin{bmatrix} 1 \\ \Delta t^{-1} \end{bmatrix}$; $P(2/2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Таким чином, після приймання другого імпульсу сигналу, стеження відбувається без похибок і $K(3) = K(4) = \dots K(n) = 0$. Більш детальне дослідження показує, що $K(k) \rightarrow 0$ якщо $Q(k) = 0$ при $R(k) \neq 0$. Незважаючи на простоту розглянутого прикладу, слід зазначити, що у процесі розрахунків отримана система зі змінними параметрами, яку важко спроектувати іншими, наприклад частотними методами.

Поряд з рекурентними алгоритмами фільтрації, при обробці інформації за допомогою цифрових ЕОМ використовуються також і нерекурентні алгоритми. У якості прикладу розглянемо алгоритм лінійного оцінювання, який заснований на методі найменших квадратів, і порівняємо його з методом фільтра Калмана. Припустимо, що є низка відліків деякої величини $y(t_1)$, $y(t_2)$, ..., $y(t_N)$ зроблених при дії завад (див. рис. 6.4).

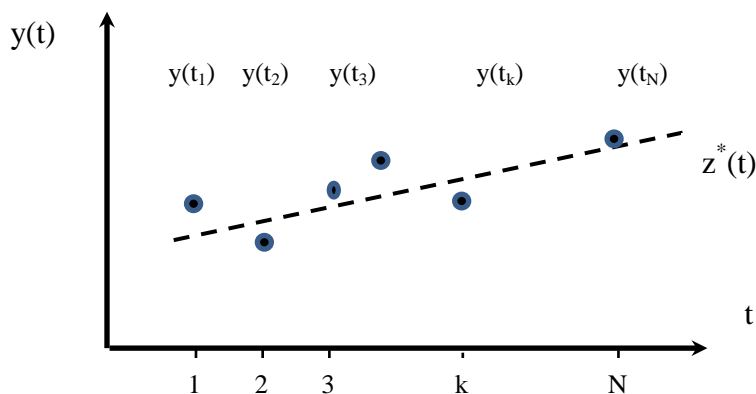


Рисунок 6.4 – Результати вимірювань з похибками

Дійсна траєкторія руху об'єкту $z(t)$ являє собою є функцію часу, що містить низку невідомих параметрів. Наприклад, зміна координат у часі може описуватися поліномом з невідомими коефіцієнтами

$$z(t) = f(t, a_i) = a_0 + a_1 t + a_2 t^2 + \dots + a_m t^m, \quad (6.15)$$

а результат вимірювання цієї величини в момент часу t_k дорівнює

$$y(t_k) = z(t_k) + v(k), \quad (6.16)$$

де $v(k)$ - дискретний "білий" шум.

У задачі, що розглядається, потрібно визначити невідомі коефіцієнти a_0, a_1, \dots, a_m за результатами низки вимірювань $y(t_1), y(t_2), \dots, y(t_N)$. Для визначення $(m+1)$ невідомих коефіцієнтів потрібна система, складена з $(m+1)$ рівняння виду (6.16). Таким чином, при $N=m+1$ можливе однозначне визначення коефіцієнтів. Однак на практиці становить інтерес розв'язок системи рівнянь при $N \gg m$, тому що при збільшенні числа відліків послабляється вплив завад $v(k)$. При $N \gg m$ однозначний розв'язок системи рівнянь не існує, проте можна знайти таку сукупність параметрів $(a_0^*, a_1^*, \dots, a_m^*)$ яка забезпечить найкращу оцінку $z^*(t_k) = f(t_k, a_i^*)$ з погляду мінімуму середнього квадрату похибки

$$\frac{1}{N} \sum_{i=1}^N [y(t_i) - z^*(t_i)]^2.$$

Векторний запис цієї задачі має вигляд:

$$Y = TA + V, \quad (6.17)$$

$$\text{де } Y = \begin{bmatrix} y(t_1) \\ \dots \\ y(t_N) \end{bmatrix}; \quad A = \begin{bmatrix} a_0 \\ a_1 \\ \dots \\ a_m \end{bmatrix}; \quad V = \begin{bmatrix} v(t_1) \\ \dots \\ v(t_N) \end{bmatrix}; \quad T = \begin{bmatrix} 1 & t_1 & \dots & t_1^m \\ 1 & t_2 & \dots & t_2^m \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_N & \dots & t_N^m \end{bmatrix}$$

Вектор похибок фільтрації дорівнює $E=Y-Z^* = Y-TA^*$

$$\text{де } A^* = \begin{bmatrix} a_0^* \\ a_1^* \\ \dots \\ a_m^* \end{bmatrix}; \quad Z^* = \begin{bmatrix} z^*(t_1) \\ z^*(t_2) \\ \dots \\ z^*(t_N) \end{bmatrix}.$$

Тоді сума квадратів похибок фільтрації становить

$$\sum_{i=1}^N [y(t_i) - z^*(t_i)]^2 = \mathbf{E}^T \mathbf{E}.$$

Знайдемо таке значення оцінки \mathbf{A}^* , при яким величина $\mathbf{E}^T \mathbf{E}$ мінімальна. Для цього обчислимо $\text{grad } \mathbf{E}^T \mathbf{E}$ і знайдемо значення \mathbf{A}^* , при якому градієнт буде дорівнювати нулю.

$$\text{grad}_{\mathbf{A}^*} \mathbf{E}^T \mathbf{E} = \text{grad}_{\mathbf{A}^*} \{[\mathbf{T}\mathbf{A}^* - \mathbf{Y}]^T [\mathbf{T}\mathbf{A}^* - \mathbf{Y}]\} = \mathbf{0}$$

Виконавши обчислення градієнту, отримаємо умову

$$2\mathbf{T}^T [\mathbf{T}\mathbf{A}^* - \mathbf{Y}] = \mathbf{0},$$

і при неособливій матриці матимемо

$$\mathbf{A}^* = [\mathbf{T}^T \mathbf{T}]^{-1} \mathbf{T}^T \mathbf{Y}. \quad (6.18)$$

У виразі (6.18) матриця $[\mathbf{T}^T \mathbf{T}]^{-1} \mathbf{T}^T$ обчислюється заздалегідь і зберігається у запам'ятовувальному пристрої. Вона визначає закон руху об'єкту й не залежить від апріорних даних про завади. Відсутність вимог до апріорних даних є істотною перевагою методу найменших квадратів.

Оцінка \mathbf{A}^* , за допомогою формули (6.18) будується шляхом обробки усього масиву даних $y(t_1), y(t_2), \dots, y(t_N)$. Така процедура вимагає значно більшого об'єму пам'яті в порівнянні з рекурентною формулою (6.14) для фільтра Калмана. Проте, виявляється, що і у випадку методу найменших квадратів на підставі виразу (6.18) можна побудувати рекурентну процедуру, у якій оцінка вектора $\mathbf{A}^*(k+1)$ після приймання $(k+1)$ -го сигналу обчислюється, виходячи з попереднього значення $\mathbf{A}^*(k)$, і нового виміру $y(k+1)$. Порівняння рекурентного виразу для методу найменших квадратів і фільтра Калмана показує, що метод найменших квадратів може розглядатися як граничний випадок фільтра Калмана при $\mathbf{P}(0)=[\infty]$ і $\mathbf{Q}(k)=0$. Якщо матриця дисперсій початкового стану заповнена нескінченними елементами, то це означає, що початковий стан лежить у нескінченних межах, а апріорна інформація про нього повністю відсутня. Рівність $\mathbf{Q}(k)=0$ приводить до того, що випадковий процес $\mathbf{x}(k)$ визначається лише випадковим вектором початкового стану $\mathbf{x}(0)$, що містить n випадкових чисел. Таким чином, вектор $\mathbf{x}(k)$ може розглядатися як функція часу k , що містить випадкові коефіцієнти $x_i(0), i=1,2,\dots,n$.

При виконанні умов $\mathbf{P}(0)=[\infty]$ і $\mathbf{Q}(k)=0$ виявляється, що значення матричного коефіцієнта підсилення $\mathbf{K}(k+1)$ у рекурентній формулі залежать тільки від номера кроку k і не залежить від $\mathbf{R}(k)$. На підставі викладеного, стає зрозумілим, чому метод найменших квадратів не вимагає використання апріорних даних $\mathbf{P}(0), \mathbf{Q}(k)$ і $\mathbf{R}(k)$.

Приклад 6.2

Розглянемо обчислення коефіцієнту підсилення у рекурентному варіанті методу найменших квадратів для найпростішого прикладу. Шуканий параметр - постійна, але невідома випадкова величина a_0 , яка виміряється з випадковою похибкою σ^2 . Обчислимо дисперсію похибки екстраполяції за формулою (6.13):

$$P(1/0) = P(0) + Q(1) = P(0) = \sigma_0^2.$$

Коефіцієнт підсилення по формулі (6.13) дорівнює

$$K(1) = P(1/0)K^T [H P(1/0) H^T + R]^{-1} = \sigma_0^2 (\sigma_0^2 + \sigma^2)^{-1}.$$

Якщо априорні дані відсутні, то $\sigma_0^2 \rightarrow \infty$, і $K(1) = I$;

$$P(1/1) = [1 - K(1)] P(1/0) = \frac{\sigma_0^2 \sigma^2}{\sigma_0^2 + \sigma^2} \rightarrow \sigma^2.$$

На другому кроці $P(2/1) = \sigma^2$, $K(2) = 1/2$ і т.д. за правилом $K(k) = 1/k$.

На цьому прикладі видно, що $K(k) \rightarrow 0$ при $k \rightarrow \infty$. Це є наслідком того, що зміна параметра відсутня і $Q(k) = 0$. Якщо ж насправді параметр повільно змінюється, то при $K(k) \rightarrow 0$ можуть виникати великі помилки, отже, на практиці слід обмежувати значення k .

7 НЕЛІНІЙНА ОПТИМАЛЬНА ФІЛЬТРАЦІЯ

7.1 Постановка задачі

Основи теорії нелінійної оптимальної фільтрації викладені у роботах [11, 17]. Для більш глибокого вивчення рекомендується [6]. В радіотехнічних системах вимірювальні параметри кодуються нелінійним чином, тому теорія лінійної фільтрації не може бути використана безпосередньо для проектуванні засобів обробки радіосигналів.

Розглянемо задачу нелінійної фільтрації випадкового процесу $\mathbf{x}(k)$, заданого векторними рівняннями

$$\mathbf{x}(k + 1) = \Phi(k)\mathbf{x}(k) + \mathbf{w}(k), \quad (7.1)$$

$$\mathbf{y}(k) = \mathbf{S}[\mathbf{H}(k)\mathbf{x}(k)] + \mathbf{v}(k). \quad (7.2)$$

Зміна параметра $\mathbf{x}(k)$, як і раніше, описується лінійним виразом, а спостережувана величина нелінійно залежить від $\mathbf{H}(k)\mathbf{x}(k)$. Шуми $\mathbf{w}(k)$ і $\mathbf{v}(k)$ являють собою векторні гаусові “білі” шуми з відомими матрицями $\mathbf{Q}(k)$ і $\mathbf{R}(k)$ та нульовими середніми значеннями. Для вирішення задачі нелінійної фільтрації за гаусової статистики оцінюваної величини $\mathbf{x}(k)$ доцільно застосовувати метод максимуму апостеріорної ймовірності [3]. Оцінювання виконується за допомогою рекурентного методу, і на першому кроці слід побудувати апостеріорний розподіл ймовірності, виходячи з апіорного розподілу ймовірності початкового стану $\mathbf{x}(0)$ та функції правдоподібності прийнятого сигналу $\mathbf{y}(t)$. Далі, на наступних кроках, у якості апіорного розподілу використовуються апостеріорні розподіли, отримані на минулих етапах.

7.2 Нелінійне оцінювання за методом максимуму апостеріорної ймовірності

Припустимо, що відома оцінка $\mathbf{x}^*(k)$, її розподіл $\mathbf{W}[\mathbf{x}(k)]$ і на основі $(k+1)$ -го вимірювання $\mathbf{y}(k+1)$ потрібно побудувати оцінку $\mathbf{x}^*(k+1)$. Складемо вираз для апостеріорного розподілу $\mathbf{W}[\mathbf{x}(k+1)/\mathbf{y}(k+1)]$ за правилом Байєса:

$$\mathbf{W}[\mathbf{x}(k+1)/\mathbf{y}(k+1)] = \frac{\mathbf{W}[\mathbf{x}(k+1)] \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]}{\mathbf{W}[\mathbf{y}(k+1)]}, \quad (7.3)$$

де $\mathbf{W}[\mathbf{x}(k+1)]$ – апіорний розподіл; $\mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]$ – функція правдоподібності; $\mathbf{W}[\mathbf{y}(k+1)]$ – розподіл сигналу $\mathbf{y}(k+1)$. У якості апіорного розподілу $\mathbf{W}[\mathbf{x}(k+1)]$ використовується розподіл оцінки $\mathbf{x}^*(k)$, екстрапольованої на один крок уперед. Припустимо, що апіорний розподіл являє собою багатовимірний нормальний розподіл з середнім значенням $\mathbf{x}^*(k+1/k)$ та кореляційною матрицею $\mathbf{P}(k+1/k)$:

$$\mathbf{W}[\mathbf{x}^*(k+1/k)] = \frac{1}{(2\pi)^{0,5n} [\det \mathbf{P}(k+1/k)]^{0,5}} \times$$

$$\times \exp\{-0,5[\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)]^T \mathbf{P}^{-1}(k+1/k)[\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)]\} \quad (7.4)$$

Величина $\mathbf{W}[\mathbf{y}(k+1)]$ також являє собою багатовимірний розподіл, який безпосередньо від \mathbf{x} не залежить. Значення $\mathbf{W}[\mathbf{y}(k+1)]$ виконують роль нормуючого множника і не впливають на розташування максимуму апостеріорного розподілу $\mathbf{W}[\mathbf{x}(k+1)/\mathbf{y}(k+1)]$. Тому надалі ця величина враховується у вигляді нормуючого множника.

Функція правдоподібності $\mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]$, у загальному випадку, має складну форму, але при високій точності вимірювань ця функція має гострий пік. Вершину цього піку з прийнятною точністю можна апроксимувати нормальним розподілом. Для виконання такої апроксимації представимо функцію правдоподібності у вигляді:

$$\mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)] = \exp\{\ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]\}.$$

Тепер проведемо розвинення цієї функції у степеневий ряд відносно точки екстраполяції $\mathbf{x}^*(k+1/k)$:

$$\begin{aligned} \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)] &= \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}^*(k+1/k)] + \\ &+ \sum_{i=1}^n \frac{\partial \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]}{\partial y_i} \Big|_{\mathbf{x}=\mathbf{x}^*(k+1/k)} [y_i(k+1) - y_i^*(k+1/k)] + \\ &+ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]}{\partial y_i \partial y_j} \Big|_{\mathbf{x}=\mathbf{x}^*(k+1/k)} \cdot \\ &[y_i(k+1) - y_i^*(k+1/k)][y_j(k+1) - y_j^*(k+1/k)] + \dots \end{aligned} \quad (7.5)$$

де $y_i^*(k+1/k) = \mathbf{H}(k+1)\mathbf{x}^*(k+1/k)$ – екстрапольований результат вимірювань на $(k+1)$ -й момент часу, а $\mathbf{y}(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1)$.

Для скорочення записів введемо такі позначення:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}; \quad a_{ij} = \frac{\partial^2 \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]}{\partial y_i \partial y_j} \Big|_{\mathbf{x}=\mathbf{x}^*(k+1/k)};$$

$$\mathbf{z} = [z_1, z_2, \dots, z_m]; \quad z_i = \frac{\partial \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)]}{\partial y_i} \Big|_{\mathbf{x}=\mathbf{x}^*(k+1/k)}.$$

В матричній формі вираз (7.5) прийме вигляд:

$$\begin{aligned} \ln W[\mathbf{y}(k+1)/\mathbf{x}(k+1)] &= \ln W[\mathbf{y}(k+1)/\mathbf{x}^*(k+1/k)]|_{\mathbf{x}=\mathbf{x}^*(k+1/k)} + \\ &+ \mathbf{z} \mathbf{H}(k+1) [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)] - 0,5 [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)]^T \times \\ &\times \mathbf{H}^T(k+1) \mathbf{A} \mathbf{H}(k+1) [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)]. \end{aligned} \quad (7.6)$$

Вираз (7.6) являє собою багатовимірну параболу

$$\exp\{\ln W[\mathbf{y}(k+1)/\mathbf{x}(k+1)]\},$$

тому його можна розглядати як апроксимацію багатовимірного нормального розподілу багатовимірною параболою навколо максимального значення. Положення максимуму апостеріорного розподілу залежить від положення максимуму добутку $W[\mathbf{x}(k+1)] W[\mathbf{y}(k+1)/\mathbf{x}(k+1)]$. З виразів (7.4) та (7.6) отримуємо:

$$\begin{aligned} W[\mathbf{x}(k+1)] W[\mathbf{y}(k+1)/\mathbf{x}(k+1)] &= \\ &= \mathbf{K} \exp\{-0,5 [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)]^T \mathbf{P}^{-1}(k+1/k) + \\ &+ \mathbf{H}^T(k+1) \mathbf{A} \mathbf{H}(k+1) [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)] + \\ &+ \mathbf{z} \mathbf{H}(k+1) [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k)], \end{aligned} \quad (7.7)$$

де \mathbf{K} – коефіцієнт, який не містить $\mathbf{x}(k)$.

Для заходження максимуму, запишемо цей добуток у вигляді:

$$\begin{aligned} &\mathbf{K}_1 \exp\left\{-\frac{1}{2} [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)]^T \times \right. \\ &\times \mathbf{P}^{-1}(k+1/k+1) [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)]. \end{aligned} \quad (7.8)$$

Такий запис зручний тим, що точка $\mathbf{x}^*(k+1/k+1)$ є точкою максимуму. Для цього у виразі (7.7) віднімемо і додамо до кожного $\mathbf{x}(k+1)$ величину $\mathbf{x}^*(k+1/k+1)$:

$$\begin{aligned} W[\mathbf{x}(k+1)] W[\mathbf{y}(k+1)/\mathbf{x}(k+1)] &= \\ &= \mathbf{K} \exp\left\{-\frac{1}{2} [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1) + \mathbf{x}^*(k+1/k+1) - \mathbf{x}^*(k+1/k)]^T \times \right. \\ &\times \mathbf{P}(k+1/k+1)^{-1} [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1) + \mathbf{x}^*(k+1/k+1) - \mathbf{x}^*(k+1/k)] + \\ &+ \mathbf{z} \mathbf{H}(k+1) [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1) + \mathbf{x}^*(k+1/k+1) - \mathbf{x}^*(k+1/k)], \end{aligned}$$

де $\mathbf{P}(k+1/k+1) = [\mathbf{P}(k+1/k)^{-1} + \mathbf{H}^T(k+1)\mathbf{A}\mathbf{H}(k+1)]$.

Слід зауважити, що $[\mathbf{x}^*(k+1/k+1) - \mathbf{x}^*(k+1/k)]$ від змінної $\mathbf{x}(k+1)$ не залежать, і вирази з цим множником можна розглядати, як постійні величини. Якщо виконати прості, проте досить громіздкі операції по зведенню подібних членів, то можна отримати вираз зведений до повного квадрату, що дає у підсумку:

$$\begin{aligned} & \mathbf{W}[\mathbf{x}(k+1)]\mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}(k+1)] = \\ & = K_1 \exp\left\{-\frac{1}{2}[\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)]^T \mathbf{P}^{-1}(k+1/k+1)[\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)] - \right. \\ & \quad - [\mathbf{x}^*(k+1/k+1) - \mathbf{x}^*(k+1/k)]^T \mathbf{P}^{-1}(k+1/k+1)[\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)] + \\ & \quad \left. + \mathbf{zH}(k+1)[\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)]\right\}. \end{aligned} \quad (7.9)$$

Порівнюючи вирази (7.8) і (7.9), можна зробити висновок, що вони співпадають за умови

$$\begin{aligned} & \{-[\mathbf{x}^*(k+1/k+1) - \mathbf{x}^*(k+1/k)]^T \mathbf{P}^{-1}(k+1/k+1) + \\ & \quad + \mathbf{zH}(k+1)\} [\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)] = 0. \end{aligned} \quad (7.10)$$

$$[\mathbf{x}(k+1) - \mathbf{x}^*(k+1/k+1)]$$

У цьому випадку, величина $\mathbf{x}^*(k+1/k+1)$ являє собою максимум апостеріорного розподілу. Таким чином, якщо виконується рівняння (7.10), то можна записати:

$$-[\mathbf{x}^*(k+1/k+1) - \mathbf{x}^*(k+1/k)]^T \mathbf{P}^{-1}(k+1/k+1) + \mathbf{zH}(k+1) = 0.$$

і далі

$$\mathbf{x}^*(k+1/k+1) = \mathbf{x}^*(k+1/k) + \mathbf{P}(k+1/k+1)\mathbf{H}^T(k+1)\mathbf{z}^T. \quad (7.11)$$

Вираз (7.11) дає структуру оптимальної оцінки. На основі попередніх вимірів обчислюється екстрапольоване значення оцінки $\mathbf{x}^*(k+1/k)$ і до нього добавляється поправка, яка залежить від прийнятого повідомлення $\mathbf{y}(k+1)$. Вага поправки визначається множником $\mathbf{P}(k+1/k+1)\mathbf{H}^T(k+1)$, який залежить від дисперсій похибок фільтрації. Тепер залишилося визначити фізичний зміст величини \mathbf{z} . Для цього візьмемо один з елементів вектора z_i та розвинемо його у ряд відносно точного значення $\mathbf{x}(k+1)$:

$$z_i = \frac{\partial}{\partial y_i} \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}]|_{\mathbf{x}=\mathbf{x}^*(k+1/k)} =$$

$$= \frac{\partial}{\partial y_i} \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}]|_{\mathbf{x}=\mathbf{x}(k+1)} + \frac{\partial^2}{\partial y_i^2} \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}]|_{\mathbf{x}=\mathbf{x}(k+1)} [y_i^*(k+1/k) - y_i^*(k+1)],$$
(7.12)

де $y_i^*(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1)$.

Для спрощення у виразі (7.12) будемо вважати, що при $i \neq j$ усі змішані похідні дорівнюють 0. Це зауваження справедливе, якщо шуми $v_i(k)$ у каналах спостереження не мають статистичного зв'язку. Пояснимо суть виразу (7.12) за допомогою графіків (рис. 7.1).

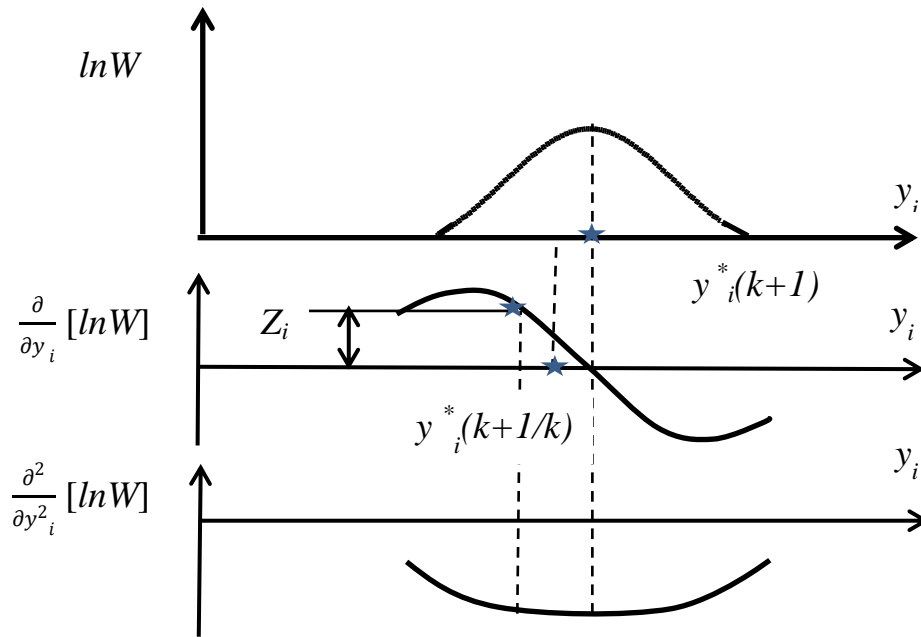


Рисунок 7.1– Ілюстрація фізичного змісту складової z_i

Логарифм функції правдоподібності має максимум поряд з точним значенням $y_i^*(k+1)$. Відхилення максимуму від істинного значення зумовлене впливом завад. Перша похідна логарифма функції правдоподібності має поблизу від точного значення $y_i^*(k+1)$ лінійну ділянку. На цій лінійній ділянці величину z_i можна подати, як суму величин:

$$\frac{\partial}{\partial y_i} \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}]|_{\mathbf{x}(k+1)} = \xi(k+1),$$

$$\frac{\partial^2}{\partial y_i^2} \ln \mathbf{W}[\mathbf{y}(k+1)/\mathbf{x}]|_{\mathbf{x}=\mathbf{x}(k+1)} [y_i^*(k+1/k) - y_i^*(k+1)] =$$

$$= K_d [y_i^*(k+1/k) - y_i^*(k+1)].$$

Величина $\xi(k+1)$ являє собою випадкову складову, тому що максимум функції правдоподібності зміщується відносно істинного значення випадковим чином. Оскільки шуми в моделі вимірювання $v(k)$ “білі”, то шум $\xi(k+1)$ також білий. Складова

$$K_d[y_i^*(k+1/k) - y_i^*(k+1)]$$

являє собою корисний сигнал похибки, зв’язаним з відхиленням оцінки від істинного значення $x(k+1)$. Таким чином, величина z_i може розглядатися, як вихідна напруга оптимального дискримінатора, що складається з суміші випадкової похибки та корисного сигналу. Крутизна характеристики дискримінатора K_d при малих похибках являє собою, практично, постійну величину. У загальному випадку, крутизна характеристики оптимального дискримінатора залежить від рівня шуму. З ростом шуму функція правдоподібності розширюється і крутизна K_d зменшується [3].

Побудуємо у відповідності до виразу (7.11) нелінійний оптимальний фільтр (рис. 7.2). Цей фільтр містить лінійну частину, яка складається з елемента затримки Δt , матриць $\Phi(k)$ і $H(k)$ і формує екстрапольоване значення оцінки $y_i^*(k+1/k)$. Нелінійні операції, що виконуються в оптимальному дискримінаторі, породжують величину z . Структура оптимального дискримінатора визначається видом функції повідомлення. Коефіцієнт передачі фільтра залежить від дисперсій похибок і дорівнює

$$P(k+1/k+1)H^T(k+1).$$

Методи нелінійної фільтрації виявляються ефективним при дослідженнях радіотехнічних систем стеження, наприклад таких, як далекоміри, кутоміри, та тощо. Метод максимуму апостеріорної ймовірності при лінійній моделі спостережень приводить до загальновідомого фільтра Калмана.

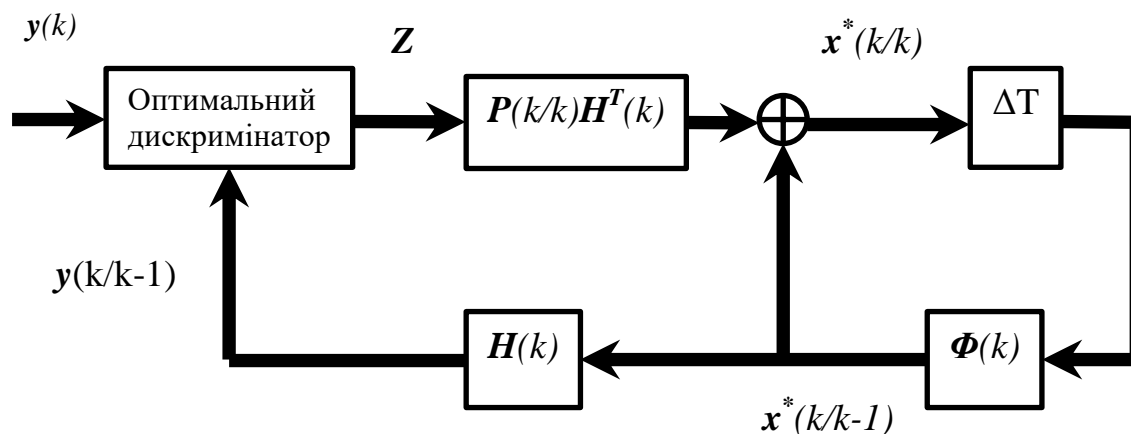


Рисунок 7.2– Структура нелінійного фільтра

Якщо спостереження описують виразом (7.8), то функція правдоподіб-

ності являє собою багатомірний гаусів розподіл і дорівнює:

$$\begin{aligned} W[\mathbf{y}(k+1)/\mathbf{x}(k+1)] &= \\ &= \frac{1}{(2\pi)^{0,5m} \det^{0,5} |\mathbf{R}(k)|} \exp\left\{-\frac{1}{2} [\mathbf{y}(k+1) - \mathbf{H}(k+1)\mathbf{x}(k+1)]^T \times \right. \\ &\quad \left. \times \mathbf{R}^{-1}(k) [\mathbf{y}(k+1) - \mathbf{H}(k+1)\mathbf{x}(k+1)]\right\}. \end{aligned} \quad (7.13)$$

Обчислимо величини \mathbf{z} та \mathbf{A} з наведеного виразу (7.13):

$$\mathbf{z} = \mathbf{R}^{-1}[\mathbf{y}(k+1) - \mathbf{H}(k+1)\mathbf{x}^*(k+1/k)]; \quad \mathbf{A} = \mathbf{R}^{-1}.$$

Таким чином, оптимальна оцінка дорівнює:

$$\begin{aligned} \mathbf{x}^*(k+1/k+1) &= \mathbf{x}^*(k+1/k) + \\ &+ \mathbf{P}(k+1)/k+1) \mathbf{H}^T(k+1) \mathbf{R}^{-1}(k+1) \times \\ &\quad \times [\mathbf{y}(k+1) - \mathbf{H}(k+1)\mathbf{x}^*(k+1/k)]. \end{aligned} \quad (7.14)$$

Покажемо, що $\mathbf{P}(k+1)/k+1) \mathbf{H}^T(k+1) \mathbf{R}^{-1}(k+1)$ співпадає з оптимальним коефіцієнтом передачі $\mathbf{K}(k+1)$. Для цього скористаємось лемою про обернення матриць, яка стверджує :

$$[\mathbf{P}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}]^{-1} = \mathbf{P} - \mathbf{P} \mathbf{H}^T (\mathbf{H} \mathbf{P} \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H} \mathbf{P}.$$

Безпосереднє застосування цієї леми завершує доказ:

$$\begin{aligned} &\mathbf{P}(k+1)/k+1) \mathbf{H}^T(k+1) \mathbf{R}^{-1}(k+1) = \\ &= [\mathbf{P}^{-1}(k+1)/k+1) + \mathbf{H}^T(k+1) \mathbf{R}^{-1}(k+1) \mathbf{H}(k+1)]^{-1} \times \\ &\quad \times \mathbf{H}^T(k+1) \mathbf{R}^{-1}(k+1) = \\ &= \mathbf{P}(k+1)/k+1) \mathbf{H}^T(k+1) [\mathbf{H}(k+1) \mathbf{P}(k+1)/k+1) \mathbf{H}^T(k+1) + \mathbf{R}(k+1)]^{-1}. \end{aligned}$$

У заключній частині розділу, розглянемо приклад застосування нелінійної фільтрації у практичній задачі.

Приклад 7.1

Задана модель зміни віддалі $d(k)$ та швидкості $v(k)$

$$\begin{bmatrix} d(k+1) \\ v(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d(k) \\ v(k) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix},$$

з відомими значеннями $\mathbf{x}(0)$, $\mathbf{P}(0)$, $\mathbf{Q}(k)$. Вимірювання віддалі виконується радіолокаційним способом. Приймається такий тип сигналу:

$$y(t, k) = U\left(t - \frac{d(k)}{0,5C}\right) \cos(\omega t) + n(t),$$

де U – закон модуляції імпульсу сигналу (рис. 7.3, а); ω – носійна частота; $n(t)$ – «білий» шум, який має задану спектральну густину N . Прийнятий сигнал має затримку, яка пов'язана з віддалю $d(k)$ і яку потрібно вимірювати оптимальним чином. При імпульсно-му випромінюванні сигналу вимірювання віддалі стає дискретним, а дискретний час k відповідає k -ому імпульсному сигналу. Для «білого» шуму функція розподілу ймовірностей дорівнює

$$W[n(t)] = K \exp\left\{-\frac{1}{N} \int_0^T n^2(t) dt\right\},$$

де $(0, T)$ – інтервал спостереження.

Для випадку, який розглядається, функція правдоподібності визначається виразом

$$W[y(t, k/d)] = K \exp\left\{-\frac{1}{N} \int_0^{T_n} [y(t, k) - u\left(t - \frac{2d}{c}\right)]^2 dt\right\},$$

де T_n – період повтору імпульсного сигналу.

Обчислимо похідну від логарифму функції правдоподібності у точці $d^*(k + 1/k)$:

$$\begin{aligned} z &= \frac{1}{N} \int_0^{T_n} 2 \left[y(t, k) - u\left(t - \frac{2d}{c}\right) \right] \frac{d}{dt} u\left(t - \frac{2d}{c}\right) dt \Bigg|_{d=d^*(k+1/k)} = \\ &= \frac{2}{N} \int_0^{T_n} u\left(t - \frac{2d^*(k+1/k)}{c}\right) \left[y(t, k) - u\left(t - \frac{2d^*(k+1/k)}{c}\right) \right] \frac{d}{dt} dt. \end{aligned}$$

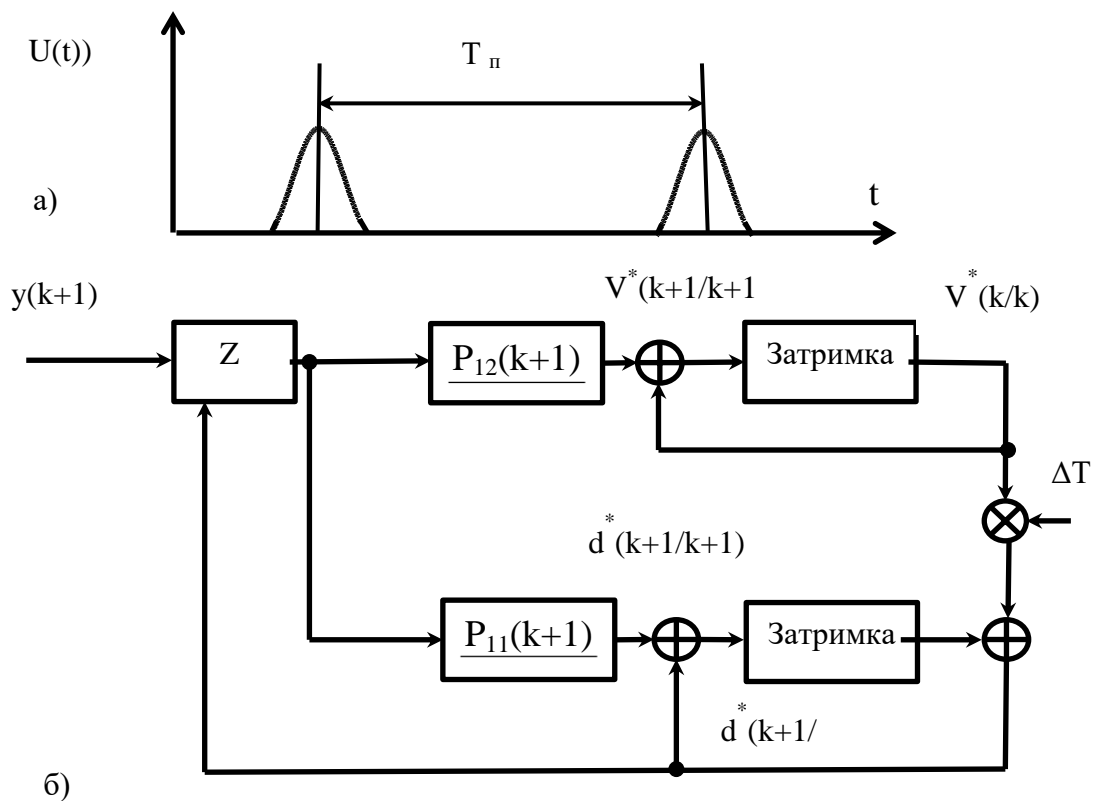


Рисунок 7.3– Структура нелінійного фільтра для прикладу 7.1

Якщо u являє собою симетричну функцію, то площа добутку

$$\left[\frac{d}{dt} u \left(t - \frac{2d}{c} \right) \right] \times u \left(t - \frac{2d}{c} \right) \quad \text{дорівнює нулю, і}$$

$$z = \frac{2}{N} \int_0^{T_n} \frac{d}{dt} \left[u \left(t - \frac{2d^*(k+1/k)}{c} \right) \right] y(t, k) dt.$$

Таким чином, у оптимальному дискримінаторі прийнятий сигнал $y(t, k)$ множиться на похідну $\frac{d}{dt} \left[u \left(t - \frac{2d^*(k+1/k)}{c} \right) \right]$, часове положення якої залежить від екстрапольованої оцінки $d^*(k+1/k)$. Результат множення інтегрується і дає значення сигналу похибки. Структура оптимального фільтру відрізняється від лінійного випадку тим, що пристрій віднімання замінено оптимальним дискримінатором, який і виконує нелінійну обробку сигналу (рис. 7.3, б). Практична реалізація оптимальних фільтрів може виконуватися двома шляхами – апаратним або алгоритмічним. У першому випадку фільтр реалізується у вигляді системи стеження, як правило – цифрової. У другому випадку алгоритми фільтрації програмуються у цифровій спеціалізованій обчислювальній машині.

Розглянемо приклад апаратної реалізації далекоміру. Відстежувальний далекомір (рис. 7.4) складається з цифрового дискримінатора, інтеграторів швидкості та віддалі, пристроїв корегування, суматора та схеми цифрової затримки. Операція оптимального розрізнення виконується у цифровому дискримінаторі, котрий виробляє послідовність імпульсів, число яких пропорційне сигналу похибки. Одночасно дискримінатор формує знак сигналу похибки. Імпульси сигналу похибки через коригувальні подільники потрапляють до лічильників інтеграторів швидкості та віддалі. Коефіцієнти поділу лічильників встановлюються у відповідності до значень оптимального коефіцієнту підсилення та можуть змінюватися у часі. Екстраполяція значень віддалі виконується шляхом періодичної добавки числа до лічильника інтегратора віддалі, яке відповідає значенню швидкості. Екстрапольоване значення оцінки віддалі надходить до схеми затримки, яка виробляє селекторний стоб-імпульс для дискримінатора. У дискримінаторі положення імпульсного сигналу приймача порівнюється з часовим положенням строб-імпульсу і потім формується напруга сигналу похибки, яка надалі перетворюється у цифрову форму.

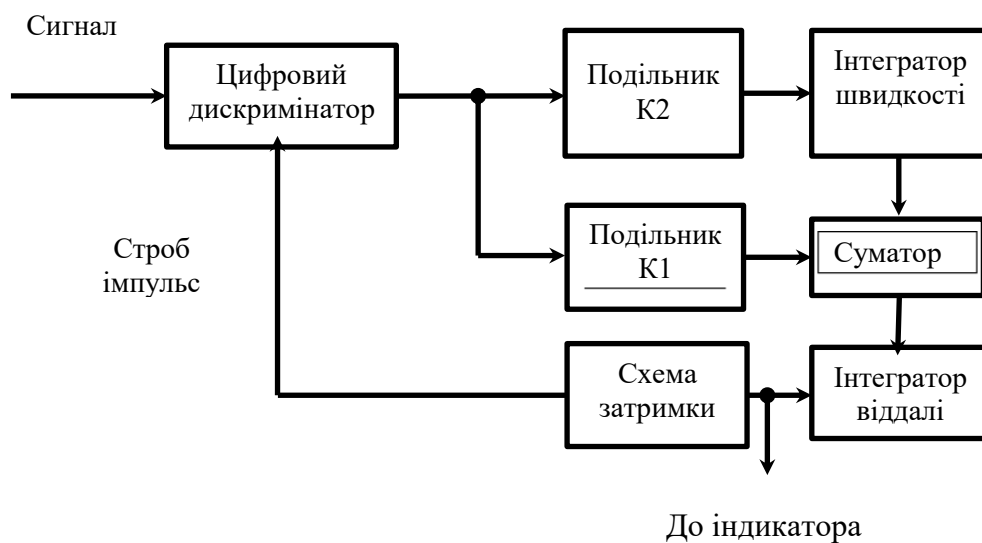


Рисунок 7.4 – Приклад апаратної реалізації цифрового далекоміра у режимі стеження

ПІСЛЯМОВА

Як міг переконатися читач, увесь вміст даного навчального посібника підкреслює необхідність всебічного розгляду цілої сукупності проблем, які виникають у процесі проектування радіотехнічних систем різного цільового призначення, а тим паче при проектуванні радіотехнічних комплексів. При використанні системного підходу проектувальник повинен мати чітку уяву про призначення системи та умов її експлуатації, тому що саме вони визначають вибір принципу роботи, тактико-технічні характеристики та структуру системи. Раціональний вибір принципу роботи та структури системи не може бути обґрунтовано вибраний без глибоких знань існуючої елементної бази та перспектив її розвитку. Проектувальник повинен уважно стежити за новітніми досягненнями у галузі теорії аналізу та синтезу сучасних радіотехнічних систем і ураховувати їх у процесі розв'язку конкретних радіотехнічних задач. Використання мікропроцесорної техніки, оптичних методів обробки сигналів, а також засобів функціональної мікроелектроніки дозволило суттєво покращити технічні характеристики радіотехнічних систем.

Радіоінженер повинен уміло застосовувати методи оптимізації складних систем, ні в якому разі не нехтувати можливостями евристичних методів оцінювання та фізичного експерименту, широко застосовувати методи математичного моделювання та систем автоматизованого проектування. На усіх етапах проектування пильно стежити за вимогами надійності та економічності виробництва і експлуатації. Оскільки у сучасній радіотехніці арсенал доступних методів та засобів безперервно поповнюється за рахунок останніх досягнень у багатьох суміжних галузях науки і техніки, то це формує надзвичайно високі вимоги до освітнянського рівня сучасного радіоінженера. У цьому сенсі дисципліна «Радіотехнічні системи» відіграє особливу роль, об'єднуючи в єдину систему знання, отримані студентом на протязі усього процесу навчання.

ПЕРЕЛІК ПОСИЛАНЬ

1. Баскаков С. И. Радиотехнические цепи и сигналы / Баскаков С. И — М. : Высшая школа, 2005. — 464 с.
2. Ширман Я. Д. Радиоэлектронные системы. Основы построения и теория / Ширман Я. Д. — М. : Радиотехника, 2007. — 512 с.
3. Радиотехнические системы : учебник для студ. высш. учеб. заведений / [Ю. М. Казаринов, Ю. А. Коломенский, Ю. П. Гришин и др.]. ; под ред. Ю. М. Казаринова. — М. : Издательский центр «Академия», 2008. С 592 с.
4. Гуткин Л. С. Проектирование радиосистем и радиоустройств / Гуткин Л. С. — М. : Радио и связь, 1986. — 472 с.
5. Трухин М. П. Математическое моделирование радиотехнических устройств и систем / Трухин М. П. — Екатеринбург: Изд-во Уральского ун-та, 2014. — 192 с.
6. Информационные технологии в радиотехнических системах: учеб. пособие / [В. А. Васин, И. Б. Власов, Д. Д. Дмитриев и др.]. ; под ред. И. Б. Федорова. — [3-е изд.]. — М. : Изд-во МГТУ им. И. Э. Баумана, 2004. — 774 с.
7. Гинзбург В. М. Голография. Методы и аппаратура / В. М. Гинзбург, Б. М. Степанов. — М. : Сов. Радио, 1974. — 376 с.
8. Перунов Ю. М. Зарубежные радиоэлектронные средства. Кн.4. Элементная база / Ю. М. Перунов, В. В. Мацукевич, А. А. Васильев. — М. : Радиотехника, 2010. — 400 с.
9. Анри Анго. Математика для электро- и радиоинженеров / Анри Анго . — М. : Наука, 1967. — 780 с.
10. Чердынцев В. А. Радиотехнические системы / Чердынцев В. А. — Минск : Вышэйшая школа, 1988. — 369 с.
11. Тихонов В. И. Нелинейная фильтрация и квазикогерентный прием сигналов / В. И. Тихонов, Н. К. Кульман. — М. : Сов. Радио, 1975. — 704 с.
12. Никольский Б. А. Основы радиотехнических систем / Никольский Б. А. — Самара : Изд-во Самар. гос. аэрокосм. ун-та, 2013. — 469 с.
13. Браммер Ю. А. Радиотехнические устройства и элементы радиосистем / Ю. А. Браммер, В. А. Каплун, С. П. Лохова — М. : Высшая школа, 2005. — 296 с.
14. Основные направления теории проектирования радиосистем / Гуткин Л. С. // Радиотехника. — 1976. — №5. — С. 2—6.
15. Гришин Ю. П. Микропроцессоры в радиотехнических системах / Ю. П. Гришин, Ю. М. Казаринов, В. М. Катиков. — М. : Радио и Связь, 1982. — 280 с.
16. Волощук Ю. І. Сигнали та процеси в радіотехніці / Волощук Ю. І. — Х. : Компанія СМІТ, 2003. — 580 с.

17. Тихонов В. И. Статистический анализ и синтез радиотехнических устройств и систем: [уч. пособие для вузов.] /Тихонов В. И., Харисов В. Н. — М. : Радио и связь, 1991. — 608 с.
18. Гуд Г. Х. Системотехника: / Г. Х. Гуд , Р. Э. Макол. Пер. с англ. — М. : Сов. Радио, 1962. — 383 с.
19. Заде Л. Теория линейных систем. Метод пространства состояний / Заде Л., Дезоер Ч. Пер. с англ. — М. : Наука, 1970 . — 704 с.
20. Бакулев П. А. Радионавигационные системы / Бакулев П. А., Сосновский А. А. — М. : Радиотехника, 2005. — 224 с.
21. Мельников Ю. П. Радиотехническая разведка. Методы оценки эффективности местоопределения источников излучения / Мельников Ю. П, Попов С. В. — М. : Радиотехника, 2008. — 432 с.

Навчальне видання

**Кичак Василь Мартинович
Воловик Андрій Юрійович
Шутило Микола Артемович
Червак Оксана Петрівна**

**РАДІОТЕХНІЧНІ СИСТЕМИ
(ОСНОВИ ПРОЕКТУВАННЯ. ЧАСТИНА 2)**

Навчальний посібник

Редактор В. Дружиніна

Оригінал-макет підготовлено А. Воловик

Підписано до друку

Формат 29,7×42¼. Папір офсетний.
Гарнітура Times New Roman.
Друк різнографічний. Ум. друк. арк.
Наклад пр. Зам. № 2018-

Вінницький національний технічний університет,
навчально-методичний відділ ВНТУ.
21021, м. Вінниця, Хмельницьке шосе, 95.
ВНТУ, к. 2201.
Тел. (0432) 59-87-36.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.

Віддруковано у Вінницькому національному технічному університеті
в комп'ютерному інформаційно-видавничому центрі
21021, м. Вінниця, Хмельницьке шосе, 95.
ВНТУ, ГНК, к. 114.
Тел. (0432) 59-87-38,
publish.vntu.edu.ua; email: kivc.vntu@gmail.com.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.