

МІНІСТЕРСТВО ОСВІТИ, НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ

Тернопільський національний технічний університет імені Івана Пулюя

Б.І.Яворський

МАТЕМАТИЧНІ ОСНОВИ РАДІОЕЛЕКТРОНІКИ

Частина 3

*Рекомендовано Міністерством освіти і науки України
як навчальний посібник*

*з дисциплін "Основи радіоелектроніки", „Фізико-теоретичні основи
конструювання радіоелектронних апаратів”, „Основи побудови та
застосування біомедичних апаратів та систем”*

Тернопіль — 2011

Яворський Б.І. Математичні основи радіоелектроніки. Частина 3. Навчальний посібник — Тернопіль: ТНТУ, 2011. — 143 с.

В третій частині посібника "Математичні основи радіоелектроніки" розглядаються методи розв'язування при допомозі радіоелектронних засобів деяких задач, що виникають в практиці при умові використання імовірнісних моделей (випадкової величини, випадкового процесу) для опису фізичних величин чи сигналу, властивих явищам, об'єктам відповідної області (медицині, геофізиці, біології, тощо).

Посібник корисний при фундаментальному початковому вивченні основ спеціальних дисциплін, що вивчаються студентами спеціальностей з напрямів базової вищої освіти (бакалавратів) "Радіоелектронні апарати", "Радіотехніка".

Іл. 18. Бібліогр.: 12 назв.

Рецензенти:

Я.П.Драган, доктор фізико-математичних наук, професор.

ЗМІСТ

Вступ	6
Розділ 1. Інтерпретація біосигналів у просторі випадкових подій	7
1.1. Імовірнісний простір	7
1.2. Імовірнісна модель експерименту із зліченим числом результатів	11
1.3. Система означень та теорем про борелівські класи множин. Теорема Каратеодорі про продовження міри	12
1.4. Імовірнісна модель експерименту з нескінченим числом результатів	15
1.5. Аксиоми теорії імовірностей	16
1.6. Умовна імовірність. Формула повної імовірності та формула Байєса	18
Підсумок	21
Розділ 2. Випадкові величини і функції розподілу	22
2.1. Функція розподілу імовірностей значень випадкової величини	22
2.2. Поняття про інтеграл. Інтеграли Лебега та Стільт'єса	26
2.3. Математичне сподівання і дисперсія випадкової величини. Моменти. Моменти вищих порядків	30
2.4. Характеристична функція	33
2.5. Поняття про якість оцінювання параметрів функцій розподілу	34
2.6. Сім'я випадкових величин	37
2.7. Евклідів простір випадкових величин	41
2.8. Кореляція. Геометрична інтерпретація кореляції	43
2.9. Рівняння регресії множини випадкових величин	45
2.10. Поняття про збіжність послідовностей випадкових величин ..	48
Підсумок	51

Розділ 3. Випадкові процеси	52
3.1. Моделі випадкового процесу	52
3.2. Неперервність та диференційовність випадкових процесів	57
3.3. Інтегрування випадкових процесів	59
3.4. Стаціонарний випадковий процес. Спектральна густина потужності стаціонарного процесу	61
3.5. Приклади стаціонарних випадкових процесів	65
3.6. Лінійні перетворення стаціонарних випадкових процесів	70
3.7. Спектральний розклад слабостаціонарного процесу	74
3.8. Стаціонарні випадкові послідовності	75
Підсумок	78
Розділ 4. Обробка сигналів у лінійних системах	79
4.1. Загальна схема обробки сигналів у системах	79
4.2. Функція розподілу відгуку лінійної ланки системи на випадковий сигнал. Нормалізація випадкового процесу	81
4.3. Кореляційна функція і енергетичний спектр відгуку лінійної ланки системи на стаціонарний випадковий процес	83
4.4. Вплив на білий шум інтегруючої ланки	85
4.4.1. Білий шум та вінерівський процес	85
4.4.2. Визначення основних характеристик процесу на виході RC-ланки	87
4.4.3. Визначення коефіцієнтів рівняння Фокера-Планка-Колмогорова	91

Розділ 5. Оптимальна фільтрація: фільтр Колмогорова-Вінера 96

5.1. Критерій оптимальності фільтру.	96
5.2. Блок-схема оптимального фільтру.	97
5.3. Вивід рівняння Вінера-Хопфа методами варіаційного числення.	99
5.4. Вивід рівняння Вінера-Хопфа методом ортогоналізації похибки.	103
5.5.Розв'язування рівняння Вінера-Хопфа.	105
5.6.Похибка оптимальної фільтрації форми сигналу.	106
Підсумок.	109

Розділ 6. Фільтр Калмана-Б'юсі. 110

6.1.Задача Калманівської фільтрації.	110
6.2.Критерій оптимальності фільтру.	112
6.3.Поновлюючий процес у фільтрі Калмана-Б'юсі.	113
6.4.Коефіцієнти фільтру Калмана-Б'юсі.	115
6.5.Структурна схема спостереження стану системи з мінімальною дисперсією похибки спостереження.	119
6.6.Алгоритм калманівської фільтрації.	121
6.7.Особливості фільтру Калмана-Б'юсі.	122
Підсумок.	123

Розділ 7. Аналіз обробки випадкових сигналів нелінійними ланками системи. 124

7.1.Статистична лінеаризація.	124
7.2.Застосування функціональних рядів для опису нелінійних ланок систем.	127
Підсумок.	134

Підсумки 136

Контрольні запитання	138
Типові задачі	140
Список використаної літератури	143

ВСТУП

Матеріал третьої частини посібника „Математичні основи радіоелектроніки” стосується теорії випадкових процесів у практиці застосувань радіоелектронної апаратури.

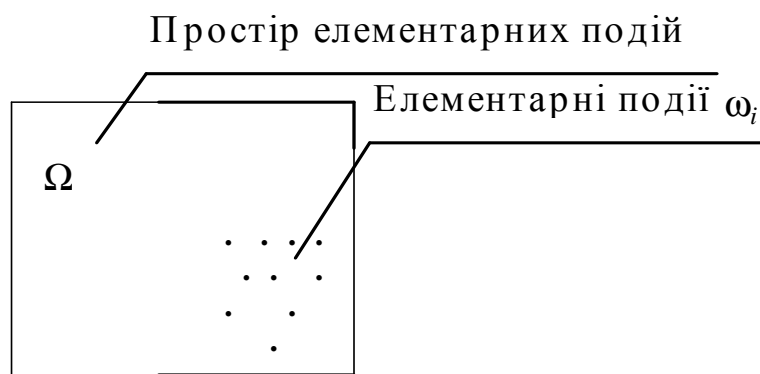
В книзі С. К. Годунова “Уравнения математической физики” задля розуміння матеріалу спеціально наголошено, що з точки зору математики, користуючись значками (частинних похідних), можна виписати самі різноманітні рівняння (в частинних похідних), але практично не всі вони застосовуються. Причому, останнє, з тієї ж математичної точки зору, є зовсім не випадковим, воно має математичне підґрунтя. Можна додати, що й фізичне (фізичну інтепретацію) також. Щось подібне стосується матеріалу, зібраного в цьому посібнику. Основним тут є не математичні сутності, а набір математичних фактів з теорії випадкових процесів, зв’язок між ними та їх тлумачення; ланцюжки: елементарна подія та її імовірність, подія і ймовірнісна алгебра, випадкова величина і функція розподілу імовірностей, параметризована сім’я випадкових величин (випадковий процес) і багатовимірна функція розподілу. Зрозуміло, що формується досить громіздка математична модель. Тому основну увагу приділено “фрагментам”, частинним, але практично важливим вислідам цієї моделі. Тут і зображення (розклади на прості взаємнонезалежні складові) складних функцій (розподілів імовірності), і використання властивостей марковості, нормальності та ін.. Коротко описати виникаючі варіанти майже неможливо. Частково відбір варіантів здійснює практика, традиції, але вони часто для розуміння, а тим більше для вивчення застосувань, теорії випадкових процесів навіть шкодять, породжують більше питань, аніж пояснюють використані математичні прийоми. Тому різні підходи до тлумачення теорії випадкових процесів у застосуваннях необхідні. В цій частині посібника описано варіант викладу, апробованого протягом кількох років.

Розділ 1

ІНТЕРПРЕТАЦІЯ БІОСИГНАЛІВ У ПРОСТОРІ ВИПАДКОВИХ ПОДІЙ

1.1. Імовірнісний простір

Елементарна подія, простір елементарних подій, випадкова подія. Розглянемо експеримент, результат якого наперед не відомий — підкидання кубика. Кубик має шість граней, кожній з яких присвоєно певне число очок (від 1 до 6). Результатом підкидання кубика буде випадання певного числа очок з ряду 1, 2, 3, 4, 5, 6, на верхній грані. Грані кубика рівноправні, тому випадання будь-якого числа очок з цього ряду є однаково можливим. Випадання якогось певного числа очок є *елементарною подією* (позначають через ω_i). У випадку з кубиком ми матимемо такі елементарні події: $\omega_1=1$; $\omega_2=2$; $\omega_3=3$; $\omega_4=4$; $\omega_5=5$; $\omega_6=6$. Сукупність усіх елементарних подій утворює множину, яка називається *простором елементарних подій* (позначають через Ω). В нашому випадку простір елементарних подій $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. Очевидно, що $\omega_i \in \Omega$ для всіх ω_i (мал.1.1).

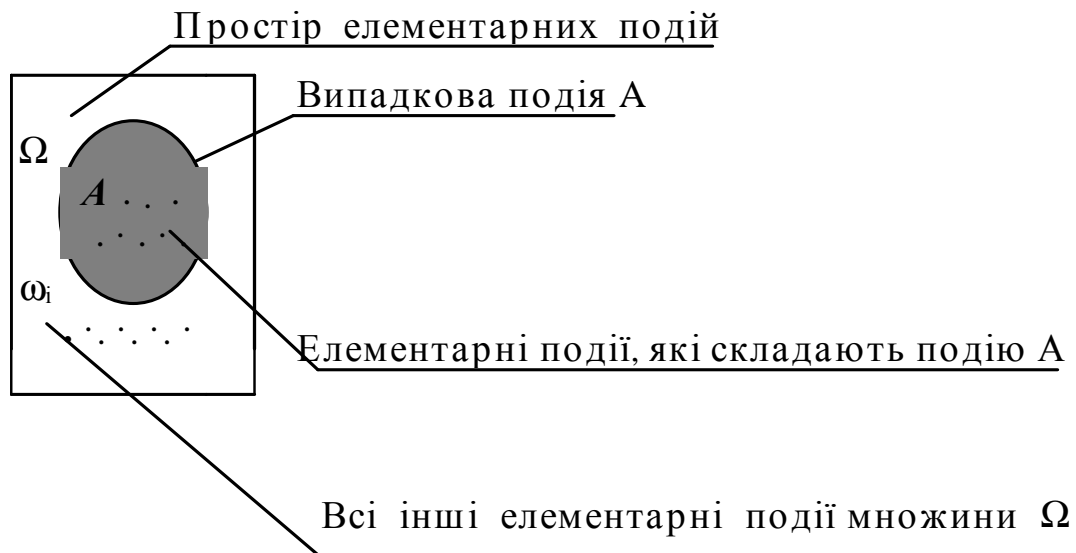


Мал. 1.1.

Простір елементарних подій може бути як скінченним (як у випадку з кубиком), так і нескінченним (тоді $i = 1, 2, \dots, \infty$, тобто $i = 1, \infty$).

Як бачимо, елементарні події взаємовиключають одна одну. Проте, в реальному досліді крім елементарних можна зауважити і складні випадкові події. В досліді з кубиком можна говорити, наприклад, про випадкову подію, яка полягає в тому, що випало парне число очок. Ця подія відбувається лише в тому випадку, коли

відбувається одна з трьох елементарних подій: ω_2 , ω_4 або ω_6 . Випаданню непарного числа очок відповідають елементарні події ω_1 , ω_3 , ω_5 . Кожну таку реальну подію можна розглядати як деяку підмножину A множини Ω , включивши в A ті і тільки ті елементарні події, при яких відбувається дана реальна подія (мал.1.2).



Мал. 1.2.

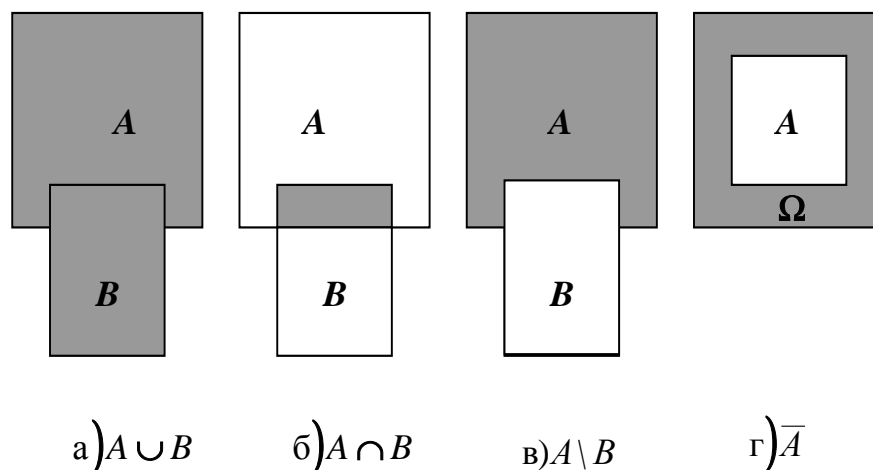
Означення 1: *Випадковою подією* (чи просто *подією*) називають будь-яку підмножину множини елементарних подій Ω , якщо Ω скінченна або зліченна. У досліді з кубиком подія $A_1 = \{\omega_2, \omega_4, \omega_6\}$ — випадання парного числа очок, подія $A_2 = \{\omega_1, \omega_3, \omega_5\}$ — випадання непарного числа очок, де A_1 , A_2 є підмножинами множини Ω (позначають $A_1 \subset \Omega$ і $A_2 \subset \Omega$).

Операції над випадковими подіями (операції над множинами):

1) **сума** (або об'єднання) двох подій A і B — подія $A+B$ (або $A \cup B$), що складається зі всіх елементарних подій, які належать події A чи B (мал.1.3,а);

2) **добуток** (або перетин) двох подій A і B — подія AB (або $A \cap B$), що складається зі всіх елементарних подій, які належать подіям A і B (мал.1.3,б);

3) **різниця** $A \setminus B$ — подія, що складається з елементарних подій множини A , які не належать B (мал.1.3,в).



Мал. 1.3.

Подію Ω називають **достовірною**, а порожню множину \emptyset — **неможливою подією**. Подія $\bar{A} = \Omega \setminus A$ називається **протилежною** події A (мал.1.3,г). Очевидно, що Ω і \emptyset — протилежні події, тобто $\Omega = \bar{\emptyset}$ і $\emptyset = \bar{\Omega}$.

Означення 2: Події A і B — **несумісні**, якщо $A \cap B = \emptyset$ (тобто множини A і B не перетинаються).

Імовірність. Для того, щоб порівнювати між собою різні події, кожній елементарній події ω_i **приписується** деяке число p_i , яке називається **імовірністю появи** цієї **елементарної події**. Так, якщо всі елементарні події рівноможливі, то у випадку, коли Ω — скінченний, $p_i = 1/N$, де N — кількість усіх елементарних подій. Тоді легко зауважити такі властивості імовірності p_i :

- 1) $p_i \geq 0$;
- 2) $\sum_{i=1}^N p_i = 1$.

Маючи імовірності елементарних подій, можна знайти імовірність будь-якої іншої події.

Означення 3: Імовірність $P(A)$ події A дорівнює сумі імовірностей елементарних подій, які входять в A :

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

Властивості імовірності подій:

- 1) $0 \leq P(A) \leq 1$;
- 2) $P(\Omega) = 1$;

3) якщо A і B – несумісні події ($A \subset \Omega$, $B \subset \Omega$), то
 $P(A \cup B) = P(A) + P(B)$.

Поняття алгебри, σ -алгебри. Як бачимо, імовірність є функцією від події. Якщо ми маємо яку-небудь подію, тобто підмножину множини Ω , то, за останнім означенням, ми можемо знайти імовірність цієї події. Проте, це не завжди так. Існують події (підмножини), для яких знайти імовірність таким способом неможливо. Тому, розглядають лише деякі підмножини із множини Ω , тобто, лише певний клас підмножин. Він відповідає певним характеристикам, що дозволяють знаходити імовірності описаним вище способом. Зокрема, він повинен бути замкнутим відносно операцій над подіями.

Означення 4: Клас підмножин A із множини Ω називається *алгеброю*, якщо:

- 1) $\Omega \in A$;
- 2) $A \in A \rightarrow \bar{A} \in A$;
- 3) $A \in A, B \in A \rightarrow A \cup B \in A, A \cap B \in A$.

Означення 5: Клас підмножин F із множини Ω називається *σ -алгеброю*, якщо:

- 1) $\Omega \in F$;
- 2) $A \in F \rightarrow \bar{A} \in F$;
- 3) $A_i \in F \rightarrow \bigcup_{i=1}^{\infty} A_i \in F, \bigcap_{i=1}^{\infty} A_i \in F, i = \overline{1, \infty}$.

Клас підмножин A (алгебра) замкнутий відносно скінченних сум і добутків, а клас підмножин F (σ -алгебра) — замкнутий відносно нескінченних сум і добутків. Тому, якщо простір елементарних подій Ω скінченний, то для нього існує тільки алгебра. Коли ж він нескінченний, то для нього існує і алгебра, і σ -алгебра.

Підведемо підсумки. Ми маємо:

- 1) простір елементарних подій Ω — множину усіх елементарних подій ω_i . У загальному випадку він нескінченний;
- 2) σ -алгебру F (так як Ω у загальному випадку нескінченний і потрібно розглядати саме σ -алгебру) — клас підмножин, для яких можливо означити і знайти імовірність події;
- 3) імовірність P — кількісну міру, що дозволяє порівнювати між собою різні події. Трійку (Ω, F, P) називають *імовірнісним простором*.

1.2. Імовірнісна модель експерименту із зліченням числом результатів

Розглянемо експеримент із випадковими результатами. Їх кількість скінченна або зліченна; всі вони рівноможливі і утворюють простір елементарних подій $\Omega = \{\omega_1, \dots, \omega_i, \dots\}$. Нехай кожній елементарній події ω_i поставлено у відповідність деяке число p_i , яке називається імовірністю елементарної події ω_i таке, що

- 1) $p_i \geq 0$;
- 2) $\sum_{i=1}^{\infty} p_i = 1$.

Тоді для випадкової події A , яка спостерігається в даному експерименті (тобто $A \subset \Omega$) імовірність $P(A)$ дорівнює сумі імовірностей елементарних подій, які входять в A ,

$$P(A) = \sum_{\omega_i \in A} p_i. \quad (1.2.1)$$

Означення 1: Якщо:

- 1) задано простір елементарних подій

$$\Omega = \{\omega_1, \dots, \omega_i, \dots\};$$

- 2) кожній ω_i із Ω приписана "імовірність" p_i з властивостями

$$p_i \geq 0, \quad \sum_{i=1}^{\infty} p_i = 1,$$

то двійку (Ω, P) називають **імовірнісною моделлю експерименту із зліченням числом результатів**.

Імовірність будь-якої події A , яка спостерігається в даному експерименті, можна знайти за формулою (1.2.1). Розглядом того, як обчислювати імовірності різних складних подій, якраз і займається теорія імовірностей. Але вона нічого не говорить про те, як "правильно" знаходити імовірності p_i елементарних подій і будувати імовірнісні моделі (Ω, P) .

1.3. Огляд означень та теорем про борелівські класи множин.

Теорема Каратеодорі про продовження міри

Розглянемо детальніше випадок зліченної множини результатів експерименту. Легко зауважити, що $p_i \rightarrow 0$, коли запроваджено метод приписування імовірності елементарній події як для випадку скінченної множини. Тому виникає питання: чи можна побудувати імовірнісну модель експерименту із зліченим числом результатів.

Перед тим, як розглянути побудову імовірнісної моделі такого експерименту, слід ознайомитись з деякими теоремами та означеннями.

Розглядатимемо деяку множину елементарних подій Ω і підмножини цієї множини. Раніше було вже дано означення алгебри і σ -алгебри:

Означення 1: Нехай A — клас підмножин з множини Ω . Якщо:

- 1) $\Omega \in A$;
- 2) $A \in A \rightarrow \bar{A} \in A$;
- 3) $A \in A, B \in A \rightarrow A \cup B \in A$,

то такий клас підмножин називається *алгеброю*.

Означення 2: Нехай F — клас підмножин з множини Ω . Якщо:

- 1) $\Omega \in F$;
- 2) $A \in F \rightarrow \bar{A} \in F$;
- 3) $A_i \in F \rightarrow \bigcup_{i=1}^{\infty} A_i \in F, i = \overline{1, \infty}$,

то такий клас підмножин називається *σ -алгеброю* (σ — сума).

Означення 3: Нехай K — деякий клас підмножин з множини Ω . σ -алгебра $\sigma(K)$ підмножин K називається *найменшою σ -алгеброю, що містить клас K* , якщо:

- 1) $K \subset \sigma(K)$;
- 2) для будь-якої іншої σ -алгебри $F, K \subset F \rightarrow \sigma(K) \subset F$.

Означення 4: Клас підмножин M називають *монотонним*, якщо:

- 1) для будь-якої послідовності множин $\{A_n\}$: $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$

з того, що $A_n \in M$ видно, що $\lim A_n = \bigcup_{n=1}^{\infty} A_n \in M$, або,

- 2) для будь-якої послідовності множин $\{A_n\}$: $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$ з

того, що $A_n \in M$ видно, що $\lim A_n = \bigcap_{n=1}^{\infty} A_n \in M$.

Теорема 1: Для того, щоб алгебра A була σ -алгеброю, необхідно і достатньо, щоб вона була монотонним класом.

Означення 5: Монотонний клас $M(K)$ називається *найменшим монотонним класом, що містить клас K* , якщо:

- 1) $K \subset M(K)$;
- 2) для будь-якого іншого монотонного класу M :

$$K \subset M \rightarrow M(K) \subset M.$$

Теорема 2: Нехай маємо алгебру A . Тоді:

$$\sigma(A) = M(A),$$

тобто, найменша σ -алгебра, що містить алгебру A і найменший монотонний клас, що містить A , співпадають.

Означення 6: Нехай множина елементарних подій

$$\Omega = R^1 = (-\infty; \infty)$$

(числова вісь) і K — клас всіх інтервалів вигляду $[a, b)$. Найменша σ -алгебра B , що містить клас K , називається *σ -алгеброю борелівських множин в R^1* . Розглянуті вище означення необхідні для того, щоб ввести поняття міри.

Означення 7: Нехай задано деяку алгебру підмножин A із Ω . Назвемо функцію множини $P(\cdot)$, означену на A , *кінцево-адитивною імовірнісною мірою* на A , якщо:

- 1) $P(A) \geq 0$, для кожного $A \in A$;
- 2) $P(\Omega) = 1$;
- 3) $P(A \cup B) = P(A) + P(B)$, коли $A \in A, B \in A, A \cap B = \emptyset$.

Означення 8: Функція множини $P(\cdot)$, означена на алгебрі A називається *зліченно-адитивною імовірнісною мірою* на A , якщо:

- 1) $P(A) \geq 0$, для кожного $A \in A$;
- 2) $P(\Omega) = 1$;

$$3) P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), \text{ коли } A_i \in \mathcal{A}, i = \overline{1, \infty}, A_i \cap A_j = \emptyset, (i \neq j),$$

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}.$$

Розглянемо умови, яким відповідає кінцево-адитивна імовірнісна міра, коли вона зліченно-адитивна.

Теорема 3: Нехай $P(\cdot)$ — кінцево-адитивна імовірнісна міра на \mathcal{A} . Вона буде зліченно-адитивною, якщо виконується одна з умов:

$$1) A_i \in \mathcal{A}, i = \overline{1, \infty}, A_i \cap A_j, (i \neq j), \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}, \text{ тоді } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i);$$

$$2) A_i, i = \overline{1, \infty}, \text{ неспадна послідовність } (A_i \subset A_{i+1}) \text{ і } \lim A_i = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A},$$

$$\text{тоді } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} P(A_i);$$

$$3) A_i, i = \overline{1, \infty} — \text{ неспадна послідовність } (A_i \supset A_{i+1}) \text{ і } \lim A_i = \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}, \text{ тоді } P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} P(A_i);$$

$$4) P(\cdot) — \text{ неперервна в } \emptyset, \text{ тобто коли } A_i \supset A_{i+1} \text{ і } \bigcap_{i=1}^{\infty} A_i = \emptyset, \text{ тоді } \lim_{i \rightarrow \infty} P(A_i) = 0.$$

При розгляді експерименту із нескінченним числом результатів використовують теорему Каратеодорі про продовження міри. Вона впливає із властивості неперервності зліченно-адитивної міри.

Теорема 4 (теорема Каратеодорі про продовження міри): Нехай існує зліченно-адитивна імовірнісна міра $P(\cdot)$ на алгебрі \mathcal{A} . Тоді існує єдина зліченно-адитивна імовірнісна міра $Q(\cdot)$, означена на найменшій σ -алгебрі $\mathcal{F} = \sigma(\mathcal{A})$, що містить \mathcal{A} і є продовженням міри P , тобто $Q(A) = P(A)$ для всіх $A \in \mathcal{A}$.

Це означає, що коли існує міра на “більшому” класі множин, то вона існує і на “меншому”.

1.4. Імовірнісна модель експерименту з нескінченним числом результатів

Розглянемо експеримент. На відрізок $[0,1)$ намання кидають точку. Вважаємо, що всі допустимі положення точки "однаково можливі". Простір елементарних подій у цьому випадку — відрізок $\Omega = [0,1)$. Побудуємо імовірнісну модель такого експерименту. Спробуємо, за аналогією з експериментом із кінцевим числом результатів, кожній елементарній події приписати значення імовірності p . Але елементарна подія в нашому випадку — це попадання точки на якусь точку відрізка $[0,1)$. Таких точок на відрізку $[0,1)$ нескінченна кількість і оскільки всі ω "однаково можливі", то імовірність кожної елементарної події $p = 0$. Тому розглядають імовірність попадання точки не в якусь точку на відрізку $[0,1)$, а на деякий інтервал, який належить заданому відрізку $[0,1)$. Таким чином, в експерименті з нескінченним числом результатів імовірності приписують не окремим елементарним подіям, а множинам елементарних подій. Основні вимоги при цьому:

— задання імовірностей повинно здійснюватись узгодженим чином, наприклад:

якщо p_1 — імовірність попадання в інтервал $[0, \frac{1}{2})$, а p_2 — в інтервал $[\frac{1}{2}, 1)$, то, очевидно, $p_1 = p_2 = \frac{1}{2}$;

якщо p_1 — імовірність попадання в інтервал $[0, \frac{1}{2})$, а p_2 — в інтервал $[0, \frac{1}{3})$, то $p_1 > p_2$ і т. д.;

— клас множин, яким приписані імовірності, повинен бути замкнутий відносно операцій об'єднання, перетину, переходу до доповнення, тобто, якщо приписані імовірності множинам A і B , то вони повинні бути приписані і множинам $A \cup B$, $A \cap B$, \bar{A} , \bar{B} .

Нехай імовірність попадання точки на деякий відрізок $[a,b) \subset [0,1)$ рівна $b - a$. Об'єднаємо певну кількість таких відрізків, що перетинаються, $\bigcup_{i=1}^m [a_i, b_i) \subset [0,1)$ і припишемо імовірність —

$\sum_{i=1}^m (b_i - a_i)$. Розглянемо множину A всіх інтервалів виду $[a,b)$ і кінцевих сум таких інтервалів, що не перетинаються. Нехай $B = \sigma(A)$ —

найменша σ -алгебра, яка містить клас A (в нашому випадку це σ -алгебра борелівських множин відрізка $[0,1)$). З теореми про продовження міри випливає, що існує єдина міра $P(\cdot)$ на B така, що $P([a,b)) = b - a$.

Отже, для σ -алгебри B можна означити імовірність, тому в якості випадкових подій розглядатимемо тільки борелівські множини із $[0,1)$. Введемо міру $P(A)$ — імовірність події $A \subset B$. Трійка (Ω, B, P) являє собою імовірнісну модель нашого експерименту, тобто, це — **імовірнісна модель експерименту з нескінченним числом результатів**.

1.5. Аксиоми теорії імовірностей

Розглянемо експеримент із випадковими результатами. Нехай Ω — простір елементарних подій, F — система підмножин із Ω , яка є σ -алгеброю, тобто

а) $\Omega \in F$;

б) $A \in F \rightarrow \bar{A} = \Omega \setminus A \in F$;

$$\text{в) } A_i \in F, i = \overline{1, \infty} \rightarrow \bigcup_{i=1}^{\infty} A_i \in F \quad (1.5.1)$$

Множини із F називають **випадковими подіями**. Нехай кожній випадковій події A (множині з F) поставлено у відповідність число $P(A)$ (назвемо його **імовірністю випадкової події A**), яке має такі властивості (1.5.2):

а) $P(A) \geq 0$, для усіх $A \in F$;

б) $P(\Omega) = 1$;

в) $\{A_i\}, i = \overline{1, \infty}$ — послідовність випадкових подій,

$$A_i \cap A_j = \emptyset \ (i \neq j), \text{ тоді } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i). \quad (1.5.2)$$

Твердження 1.5.1 (а-в), 1.5.2 (а-в) є аксіомами теорії імовірності (зформульовані А. М. Колмогоровим). Трійку (Ω, F, P) , де Ω — простір елементарних подій, F — σ -алгебра підмножин із Ω , а $P(\omega)$ — імовірність (або імовірнісна міра) на F , називають **імовірнісним простором**. Якщо побудований імовірнісний простір (Ω, F, P) , то

кажуть, що побудована імовірнісна модель експерименту. Побудова імовірнісного простору в кожному з випадків є складною, їй передують детальне вивчення конкретних умов, виявлення їх властивостей, перевірка.

Властивості імовірності. На основі наведених аксіом можна встановити наступні властивості імовірності:

- 1) $P(\overline{A}) = 1 - P(A)$;
 а) $P(\emptyset) = 0$;
- 2) якщо $A \subset B$, то $P(B \setminus A) = P(B) - P(A)$;
 а) якщо $A \subset B$, то $P(A) \leq P(B)$;
 б) для усіх A : $P(A) \leq 1$;
- 3) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ — **теорема додавання**;
 а) якщо $A \cap B = \emptyset$ (події несумісні), то $P(A \cup B) = P(A) + P(B)$;
- 4) якщо $\{A_n\}$ — монотонно неспадна, тобто

$$A_1 \subset A_2 \subset \dots \subset A_n \subset A_{n+1} \subset \dots,$$

то

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n) ;$$

— якщо $\{A_n\}$ — монотонно незростаюча, тобто

$$A_1 \supset A_2 \supset \dots \supset A_n \supset A_{n+1} \supset \dots,$$

то

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n) ;$$

$$4.1) P\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n) \text{ — властивість неперервності};$$

5) якщо $\{A_n\}$ — скінченна або зліченна, то

$$P\left(\bigcup_n A_n\right) \leq \sum_n P(A_n) ;$$

$$P\left(\bigcap_n A_n\right) \geq 1 - \sum_n P(A_n) .$$

1.6. Умовна імовірність. Формула повної імовірності та формула Байєса

Умовна імовірність. При вивченні реальних випадкових явищ інколи доводиться розглядати імовірність деякої випадкової події B , коли відомо, що відбулася деяка інша подія A , що має вплив на B .

Приклад: Уявімо, що студент із 30 білетів встиг вивчити білети від 1-го по 3-ій і від 28-го по 30-ий. На екзамен він прийшов одинадцятим, і виявилось, що до його приходу залишились тільки білети від 1-го до 20-го (подія A). Нехай подія B — студент витягнув вивчений білет. Імовірність цієї події без додаткової інформації про те, що подія A відбулася,

$$P(B) = \frac{|B|}{|\Omega|} = \frac{6}{30} = \frac{1}{5}$$

де $\Omega = \{1, 2, \dots, 30\}$ — можливі результати досліду (номери білетів). При додатковій інформації (подія A відбулася) множина можливих результатів події A складається з $|A|=20$ елементарних результатів, а подія B разом з A відбувається в $|AB|=3$ випадках. Отже, можна ввести умовну імовірність $P(B|A)$ події B при умові, що подія A відбулася,

$$P(A|B) = \frac{|AB|}{|A|} = \frac{3}{20}$$

Поділивши чисельник і знаменник на $|\Omega|$, одержимо:

$$P(A|B) = \frac{P(AB)}{P(A)} \quad (1.6.1)$$

Формула (1.6.1) — загальне означення умовної імовірності.

Означення 1: Нехай (Ω, \mathcal{F}, P) — довільний імовірнісний простір, A — випадкова подія, $A \in \mathcal{F}$, $P(A) > 0$. **Умовною імовірністю** події B за умови, що відбулася подія A , називають імовірність

$$P(B|A) = \frac{P(AB)}{P(A)}$$

Властивості:

- 1) $0 \leq P(B|A) \leq 1$;
- 2) $P(\Omega|A)=1$;
- 3) $P(A|A)=1$.

Із формули (1.6.1) $P(AB) = P(A)P(B|A)$. Якщо $P(B)>0$, то $P(AB) = P(B)P(A|B)$. Отже, коли $P(A) > 0$ і $P(B) > 0$, то

$$P(AB)=P(A)P(B|A)=P(B)P(A|B) \quad (1.6.2)$$

— **формула множення.**

Якщо ж $P(AB)=P(A)P(B)$, то події A і B називають **незалежними**.

Формула повної імовірності.

Означення 2: Набір випадкових подій B_1, \dots, B_n утворює **повну групу подій**, якщо:

- 1) $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$;
- 2) $B_i \cap B_j = \emptyset$ ($i \neq j$), $i = \overline{1, n}, j = \overline{1, n}$

Теорема 1: Якщо B_1, \dots, B_n утворюють повну групу подій і $P(B_i) > 0$, $i = \overline{1, n}$, то імовірність будь-якої випадкової події $A \in F$

$$P(A) = \sum_{i=1}^n P(B_i)P(A|B_i) \quad (1.6.3)$$

Рівність (1.6.3) — **формула повної імовірності**.

Формула Байєса. Згідно формули (1.6.2) умовна імовірність

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

звідки

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(A)}$$

Користуючись формулою повної імовірності (1.6.3), запишемо

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{i=1}^n P(B_i)P(A|B_i)} \quad (1.6.4)$$

Формула (1.6.4) — **формула Байєса**.

Підсумок

У цьому розділі викладено основні факти теорії імовірності для випадкової події так, щоб було видно перш за все її логічну структуру, алгебру, або, як ще кажуть, теоретико-множинну основу. Основний зміст стосується поняття імовірності як міри (випадкової). Ці дані згодом використовуватимуться і під час ознайомлення з теорією прийняття рішень, і під час вивчення методів аналізу сигналів. Але перш за все поняття імовірності необхідне у наступному розділі, коли розглядається не подія, а якась фізична величина, значення якої tworять множину Ω (тоді розумно припустити, що вони належать імовірнісному просторові). Отже, для коректного опису (моделювання) цієї величини необхідні відповідні імовірнісні підходи. Кажуть, що даній величині властива імовірнісна структура. Що собою являє ця структура, як вона допомагає сукупному розглядові двох величин, які проблеми розв'язує — таке уявлення про досліджувану величину, описано в наступних розділах.

Розділ 2

ВИПАДКОВІ ВЕЛИЧИНИ І ФУНКЦІЇ РОЗПОДІЛУ

2.1. Функція розподілу імовірностей значень випадкової величини

Розглянемо експеримент з числовим результатом. Результати експерименту можна вважати за елементарні події. Тоді результат — це множина елементарних подій. Оскільки елементарна подія є число ξ , то результат експерименту є подією, яку називають **випадковою величиною**. Отже, випадкова величина — це функція $\xi \equiv \xi(\omega)$ на просторі елементарних подій Ω , або $\varepsilon: \xi(\omega)$, $\omega \in \Omega$.

Приклад 1: Кидають кубик. В цьому випадку простір елементарних подій

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\},$$

де ω_i означає, що випало i очок. Випадкова величина ξ — кількість очок, яка випала, — є функцією елементарної події, причому $\xi(\omega_i) = i$.

Конструктивною характеристикою випадкової величини є функція, означена на Ω і така, що $\{\omega: \xi(\omega) \in (-\infty, x)\} = \{\omega: \xi(\omega) < x\}$, $x \in R$ і належить σ -алгебрі F випадкових подій.

Означення 1: Нехай (Ω, F, P) — імовірнісний простір. **Випадковою величиною** ξ називають дійсну функцію $\xi(\omega)$, таку, що при довільному дійсному x

$$\{\omega: \xi(\omega) < x\} \in F. \quad (2.1.1)$$

В теорії функцію $\xi(\omega)$, коли для кожного дійсного x виконана умова (2.1.1), називають **вимірною відносно σ -алгебри F функцією**. Випадкові величини описуються функціями розподілу імовірностей їх значень.

Означення 2: Функцією розподілу імовірностей значень випадкової величини $\xi(\omega)$ називається функція

$$F_{\xi}(x) = P\{\omega: \xi(\omega) < x\}. \quad (2.1.2)$$

де P — імовірність.

Приклад 2: Підкидається монета. В цьому випадку $\Omega=(\omega_1, \omega_2)$, де $\omega_1 \equiv \Gamma$ (випав герб), $\omega_2 \equiv \Pi$ (випала цифра), і $P(\{\omega_1\})=P(\{\omega_2\})=\frac{1}{2}$. Якщо при підкиданні випав герб, то нараховується 1 очко, а якщо випала цифра, то віднімається 1 очко, тобто випадкова величина $\xi(\omega)$ задається наступним чином:

$$\xi(\omega) = \begin{cases} 1, & \omega = \omega_1; \\ -1, & \omega = \omega_2. \end{cases}$$

Очевидно, що подія

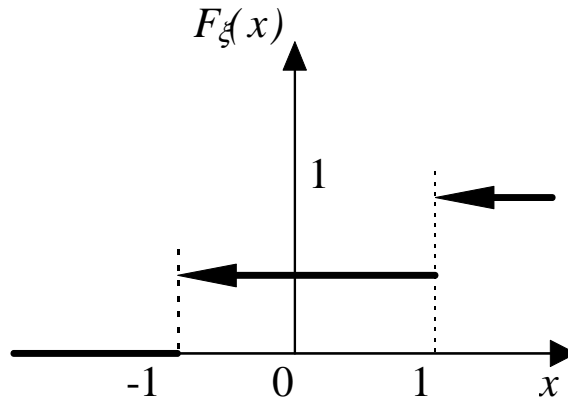
$$\{\omega: \xi(\omega) < x\} = \begin{cases} \emptyset, & x \leq -1; \\ \{\omega\}, & -1 < x \leq 1; \\ \Omega, & x > 1. \end{cases}$$

Тоді функція розподілу імовірностей значень випадкової величини

$$F_{\xi}(x) = P\{\omega: \xi(\omega) < x\} = \begin{cases} 0, & x \leq -1; \\ \frac{1}{2}, & -1 < x \leq 1; \\ 1, & x > 1. \end{cases}$$

Графік функції розподілу зображено на мал.2.1.

Функцію розподілу імовірностей значень випадкової величини $F_{\xi}(x)$ називають ще **інтегральною функцією розподілу**. Властивості інтегральної функції розподілу:



Мал. 2.1.

- 1) якщо $x_1 < x_2$, то $P\{\omega: x_1 \leq \xi(\omega) < x_2\} = F_\xi(x_2) - F_\xi(x_1)$;
- 2) якщо $x_1 < x_2$, то $F_\xi(x_1) \leq F_\xi(x_2)$, тобто $F_\xi(x)$ — неспадна функція;
- 3) $\lim_{x \rightarrow -\infty} F_\xi(x) = 0$; $\lim_{x \rightarrow +\infty} F_\xi(x) = 1$; тобто $0 \leq F_\xi(x) \leq 1$;
- 4) $\lim_{x \rightarrow x_0 - 0} F_\xi(x) = F_\xi(x_0)$, тобто $F_\xi(x)$ — неперервна зліва.

Теорема 1: Нехай $F(x)$ — деяка функція, що має властивості:

- 1) $F(x)$ — неспадна на $(-\infty; +\infty)$;
- 2) $F(x)$ — неперервна зліва;
- 3) $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow +\infty} F(x) = 1$, тоді $F(x)$ — буде функцією

розподілу деякої випадкової величини $\xi(\omega)$ на деякому імовірнісному просторі (Ω, \mathcal{F}, P) .

Якщо функція розподілу імовірностей випадкової величини неперервна, то існує функція густини розподілу імовірностей. **Густиною розподілу імовірностей значень випадкової величини ξ** називається функція $p_\xi(x)$, для якої

$$F_\xi(x) = \int_{-\infty}^x p_\xi(u) du. \quad (2.1.4)$$

Густину розподілу імовірностей значень випадкової величини називають ще диференціальною функцією розподілу. Властивості диференціальної функції розподілу:

$$1) p(x) = F'(x) \text{ в точках неперервності } p(x);$$

$$2) p(x) \geq 0;$$

$$3) P\{a \leq \xi < b\} = \int_a^b p(u) du;$$

$$4) \int_{-\infty}^{\infty} p(u) du = 1. \quad (2.1.5)$$

За аналогією із функцією розподілу, кожна невід'ємна функція $p(x)$, яка має властивість (2.1.5), буде густиною розподілу для деякої випадкової величини.

Три типи функцій розподілу. Розрізняють:

1) дискретну функцію розподілу $F(x)$ — функція розподілу дискретних випадкових величин:

$$F_D(x) = \sum_{x_i < x} p(x_i),$$

де $p(x_i)$ — імовірність;

2) абсолютно неперервну функцію розподілу $F_{AH}(x)$ — функція розподілу випадкової величини, яка має густину розподілу $p(x)$,

$$F_{AH}(x) = \int_{-\infty}^x p(u) du ;$$

3) сингулярну функцію розподілу F_C — функція розподілу величини, яка існує тільки в точці (наприклад, детермінованої величини).

Теорема 2 (теорема Лебега): Будь-яка функція розподілу може бути записана у вигляді суми

$$F(x)=F_D+F_{AH}+F_C.$$

2.2. Поняття про інтеграл. Інтеграли Лебега та Стільт'єса

Поняття про інтеграл як границю “інтегральної суми” з доданками — добутками значення функції на приріст аргументу в цій точці, пов’язують з іменем Рімана (Rieman J. V.). При цьому інтегровна функція має бути неперервною. Функції розподілів у загальному випадку розривні. Крім того, при обробці імовірнісних сигналів застосовують інтегрування. Але значення таких сигналів мають різну імовірнісну міру, диференціали їх мають різну “вагу” у інтегральних сумах. Тому, хоч би задля коректності виразів, потрібно ці факти враховувати. Існують інші способи означення інтегралу, інтегрування. Розглянемо інтеграл по мірі на множині та інтеграл по функції.

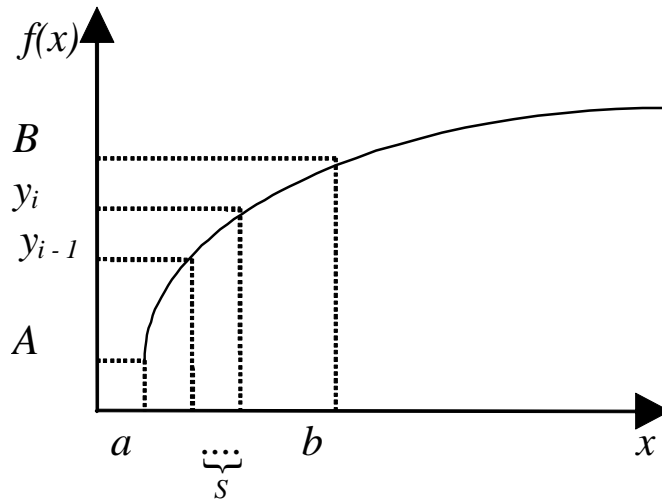
Інтеграл Лебега. Ми вже розглядали поняття міри. Міра $M(s)$, означена на певному класі множини s — це функція множини із властивостями:

$$1) M(s) \geq 0;$$

$$2) M(\emptyset) = 0;$$

$$3) M\left(\bigcup_{i=1}^{\infty} s_i\right) = \sum_{i=1}^{\infty} M(s_i), \text{ коли } s_i \cap s_j = \emptyset, (i \neq j).$$

Мірою Лебега $m(s)$ називають міру, означену на прямій, що має додаткову властивість $m(s) = m([a, b]) = b - a$ для кожного обмеженого інтервалу $[a, b]$. Множина S , для якої існує міра Лебега, називається **вимірною множиною**. Якщо функція $f(x)$ означена на інтервалі $[a, b]$ і для кожного дійсного числа c множина точок x інтервалу $[a, b]$, в яких $f(x) \leq c$, вимірна, то кажуть, що $f(x)$ — **вимірна функція** на $[a, b]$. Нехай дійсна функція $y = f(x)$ вимірна і обмежена на інтервалі $[a, b]$, A і B — відповідно її нижня і верхня границі. Розіб'ємо інтервал $[A, B]$, що містить множину значень функції $f(x)$ на $[a, b]$, на n частин (мал. 2.2.)



Мал. 2.2.

$$A = y_0 < y_1 < y_2 < \dots < y_n = B,$$

і позначимо через s_i множину точок x на інтервалі $[a, b]$, в яких

$$y_{i-1} < f(x) \leq y_i.$$

Складемо дві суми (інтегральні суми Лебега) $\sum_{i=1}^n y_{i-1} m(s_i)$ і $\sum_{i=1}^n y_i m(s_i)$.

За умови, що $\max(y_i - y_{i-1}) \rightarrow 0$, обидві інтегральні суми прямують до однієї і тієї ж границі, яка не залежить від вибору значень y_i . Число

$$I = \lim_{\max(y_i - y_{i-1}) \rightarrow 0} \sum_{i=1}^n y_i m(s_i) = \int_a^b f(x) dm = \int_S f(x) dm,$$

де S — множина точок x інтервалу $[a, b]$, називають **інтегралом Лебега** від функції $f(x)$ на інтервалі $[a, b]$ (чи на множині S).

Інтеграл Лебега $\int_S f(x) dm$ існує тоді і тільки тоді, коли існує

інтеграл Лебега $\int_a^b |f(x)| dm$. Якщо існує і абсолютно збігається

довільний інтеграл Рімана, то існує і відповідний інтеграл Лебега, який

рівний інтегралу Рімана. Інтеграл Лебега має властивості, аналогічні інтегралу Рімана:

$$1) \int_S f(x) dm = \int_{S_1} f(x) dm + \int_{S_2} f(x) dm, \text{ якщо } S=S_1 \cup S_2 ;$$

$$2) \int_S [f_1(x) + f_2(x)] dm = \int_S f_1(x) dm + \int_S f_2(x) dm ;$$

$$3) \int_S af(x) dm = a \int_S f(x) dm .$$

Перевага інтегралу Лебега полягає в тому, що його можна застосувати до більш широкого класу функцій, ніж інтеграл Рімана. Це спрощує формулювання багатьох теорем.

Інтеграл Стільт'єса. Нехай дійсна функція $f(x)$ означена і обмежена на інтервалі $[a,b]$. Розіб'ємо цей інтервал на n менших інтервалів точками $a=x_0 < x_1 < x_2 < \dots < x_n = b$.

Виберемо в кожному із таких інтервалів по довільній точці ξ_i ($\xi_i \in [x_{i-1}, x_i]$). Введемо функцію $g(x)$ з певними властивостями, яку назовемо **інтегруючою функцією**. Складемо суму (інтегральну суму)

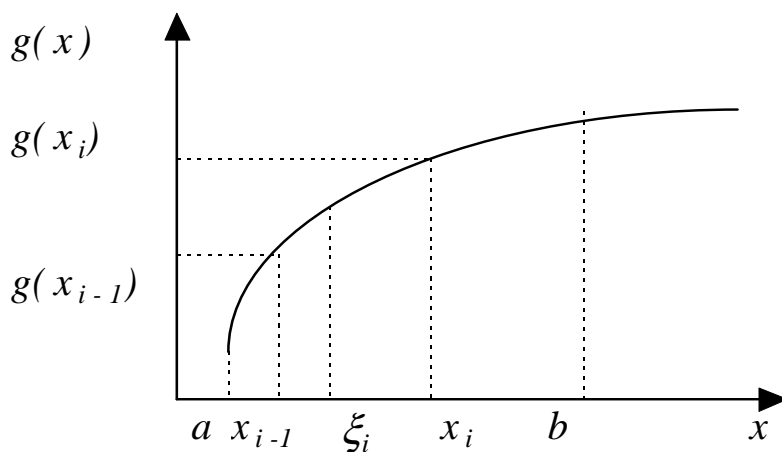
$$\sum_{i=1}^n f(\xi_i) [g(x_i) - g(x_{i-1})]$$

Інтегралом Стільт'єса від функції $f(x)$ з інтегруючою функцією $g(x)$ на інтервалі $[a,b]$ називається границя інтегральної суми за умови, що $\max(x_i - x_{i-1}) \rightarrow 0$, тобто:

$$\int_a^b f(x) dg(x) = \lim_{\max(x_i - x_{i-1}) \rightarrow 0} \sum_{i=1}^m f(\xi_i) [g(x_i) - g(x_{i-1})]$$

.

Для наочності можна розглянути простіший випадок, коли функція $f(x)=x$ (мал. 2.3).



Мал. 2.3.

Інтегральна сума в цьому випадку матиме вигляд:

$$\int_a^b x dg(x) = \lim_{\max(x_i - x_{i-1}) \rightarrow 0} \sum_{i=1}^m \xi_i [g(x_i) - g(x_{i-1})].$$

Властивості інтегралу Стільт'єса:

- 1) $\int_a^b f dg = - \int_b^a f dg;$
- 2) $\int_a^b f dg = \int_a^c f dg + \int_c^b f dg;$
- 3) $\int_a^b (f_1 + f_2) dg = \int_a^b f_1 dg + \int_a^b f_2 dg;$
- 4) $\int_a^b f d(g_1 + g_2) = \int_a^b f dg_1 + \int_a^b f dg_2;$
- 5) $\int_a^b (af) dg = \int_a^b f d(ag) = a \int_a^b f dg.$

Інтеграли Стільт'єса часто мають "наочний" зміст: ними зручно описувати криволінійні інтеграли, інтеграли по поверхні, по об'єму, по розподілу (маси, заряду, імовірності, інші) на поверхні, в об'ємі, на осі.

2.3. Математичне сподівання і дисперсія випадкової величини.

Моменти. Моменти вищих порядків

Математичне сподівання. Функція розподілу повністю визначає випадкову величину. Але часто можна обмежитись характеристиками функції розподілу чи її параметрами. Одним з важливих параметрів функції розподілу є математичне сподівання. Воно легко обчислюється: середнє арифметичне рівномірно розподілених значень випадкової величини близьке до математичного сподівання.

Означення 1: Математичним сподіванням $M\xi$ дискретної випадкової величини ξ , яка приймає значення x_1, x_2, \dots з імовірностями p_1, p_2, \dots , називається число

$$M\xi = \sum_{k=1}^{\infty} p_k x_k ,$$

якщо ряд абсолютно збігається.

Означення 2: Математичним сподіванням $M\xi$ довільної випадкової величини ξ , заданої на імовірнісному просторі (Ω, F, P) , називається число

$$M\xi(\omega) = \int_{\Omega} \xi(\omega) dP(\omega) ,$$

якщо інтеграл Лебега від функції $\xi(\omega)$ по мірі $P(\cdot)$, який стоїть в правій частині рівності, існує.

---- **Властивості математичного сподівання** (безпосередньо впливають із властивостей інтегралу Лебега):

- 1) якщо c —стала, то $Mc=c$;
- 2) якщо c —стала, то $M(c\xi)=cM\xi$;
- 3) для довільних випадкових величин ξ

$$|M\xi| \leq M|\xi|;$$

- 4) для довільних випадкових величин ξ_1 і ξ_2

$$M(\xi_1 + \xi_2) = M\xi_1 + M\xi_2;$$

5) якщо випадкові величини ξ_1 і ξ_2 незалежні, то

$$M(\xi_1 \xi_2) = M\xi_1 M\xi_2 .$$

Математичне сподівання довільної випадкової величини зручно обчислювати за її функцією розподілу. Використовуючи інтеграл Стільт'єса, можна записати формулу для обчислення $M\xi$ у вигляді:

$$M\xi = \int_{-\infty}^{+\infty} x dF_{\xi}(x)$$

де $F(x)$ — функція розподілу випадкової величини ξ .

В частинному випадку, для неперервної випадкової величини, якщо $p_{\xi}(x)$ — її густина розподілу, формула матиме вигляд:

$$M\xi = \int_{-\infty}^{+\infty} x p_{\xi}(x) dx , \quad (2.3.1)$$

Саме цією формулою користуються для знаходження математичного сподівання неперервної випадкової величини.

Моменти. Для числової характеристики випадкової величини використовують моменти функції густини розподілу. Розрізняють початкові і центральні моменти.

Означення 3: *Початковим моментом порядку k* називається число

$$m_{\xi}^k = M\xi^k = \int_{-\infty}^{+\infty} x^k p(x) dx .$$

Очевидно, що математичне сподівання є початковим моментом 1-го порядку, тобто:

$$M\xi = m_{\xi}^1 = m_{\xi} . \quad (2.3.2)$$

Означення 4: *Центральним моментом порядку k* називається число

$$\mu_{\xi}^k = M(\xi - M\xi)^k = \int_{-\infty}^{+\infty} (x - m_{\xi})^k p_{\xi}(x) dx. \quad (2.3.3)$$

Центральний момент 1-го порядку рівний нулеві. Справді, так як $M\xi$ — постійна величина, то $M(M\xi) = M\xi$ і

$$\mu_{\xi}^1 = M(\xi - M\xi) = M\xi - M\xi = 0.$$

На практиці користуються переважно початковим моментом 1-го порядку (математичним сподіванням) і центральним моментом 2-го порядку (дисперсією).

Дисперсія. Крім математичного сподівання важливим параметром густини функції розподілу випадкової величини є дисперсія.

Означення 5: *Дисперсією $D\xi$* випадкової величини ξ називається центральний момент 2-го порядку, тобто число

$$D\xi = \mu_{\xi}^2 = M(\xi - M\xi)^2. \quad (2.3.4)$$

Дисперсія характеризує “розсіювання” значень випадкової величини навколо її математичного сподівання. Величину $\sigma_{\xi} = \sqrt{D\xi}$ називають *середньоквадратичним відхиленням*.

Якщо скористатися властивостями математичного сподівання, то праву частину (2.3.4) можна звести до вигляду:

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

Отже,

$$D\xi = M\xi^2 - (M\xi)^2. \quad (2.3.5)$$

Якщо випадкова величина ξ дискретна, то

$$D\xi = \sum_k (x_k - M\xi)^2 p_k.$$

У випадку неперервної випадкової величини —

$$D\xi = \int_{-\infty}^{+\infty} (x - M\xi)^2 p_{\xi}(x) dx. \quad (2.3.6)$$

Властивості дисперсії:

- 1) для довільної випадкової величини ξ : $D\xi \geq 0$;
- 2) якщо c — стала, то $Dc=0$;
- 3) якщо c — стала, то $D(c\xi)=c^2D\xi$;
- 4) якщо випадкові величини ξ_1 і ξ_2 незалежні, то

$$D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2 .$$

2.4. Характеристична функція

При вивченні випадкових величин використовують також характеристичні функції розподілу цих випадкових величин.

Означення 1: *Характеристичною функцією* $\varphi(z)$ випадкової величини ξ називається комплекснозначна функція, означена при $z \in C$ співвідношенням

$$\varphi(z) = M e^{izx} = M (\cos(zx) + i \sin(zx)) . \quad (2.4.1)$$

Якщо $F(x)$ — функція розподілу величини ξ , то

$$\varphi(z) = \int_{-\infty}^{\infty} e^{izx} dF(x). \quad (2.4.2)$$

Характеристична функція однозначно визначає розподіл випадкової величини.

Розглянемо деякі властивості характеристичної функції:

- 1) $\varphi(0)=1$;
- 2) $|\varphi(z)| \leq 1$ для всіх дійсних z ;

$$3) \varphi(-z) = \overline{\varphi(z)}.$$

Одну із основних властивостей характеристичної функції сформулюємо у вигляді теореми.

Теорема 1: Якщо ξ_1 і ξ_2 — незалежні випадкові величини, а $\varphi_1(z)$ і $\varphi_2(z)$ — їх характеристичні функції, то характеристична функція суми $\xi_1 + \xi_2$ рівна добутку $\varphi_1(z)\varphi_2(z)$.

Як наслідок, для k незалежних випадкових величин $\xi_1, \xi_2, \dots, \xi_k$ характеристична функція суми $\xi_1 + \xi_2 + \dots + \xi_k$ рівна добутку характеристичних функцій доданків.

Теорема 2: Нехай ξ_1 і ξ_2 — незалежні нормально розподілені випадкові величини з параметрами a_1, σ_1^2 і a_2, σ_2^2 відповідно, тобто

$$F_{\xi_i}(x) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-a_i)^2}{2\sigma_i^2}}. \quad \text{Тоді } \xi_1 + \xi_2 \text{ має нормальний розподіл з}$$

параметрами $a_1 + a_2, \sigma_1^2 + \sigma_2^2$. Як видно характеристичні функції корисні при дослідженнях сум незалежних випадкових величин.

2.5. Поняття про якість оцінювання параметрів функцій розподілу

Розділ математики, який вивчає методи знаходження характеристик випадкових величин називають математичною статистикою. Іноді формули для оцінювання параметрів характеристик випадкових величин називають статистиками. У математичній статистиці має місце **задача оцінювання невідомих параметрів** наперед вибраної функції розподілу за результатами спостережень x_1, x_2, \dots, x_n — **вибіркою**. Розглядаються також випадкові вектори $(\xi_1, \xi_2, \dots, \xi_n)$, де $\xi_k, k = 1, n$ — незалежні і розподілені з деякою функцією розподілу $F_{\xi}(x)$ випадкові величини. Вибірка повинна бути репрезентативною, тобто нести сукупність ознак випадкової величини. До властивостей випадкових величин за їх значеннями (**реалізаціями**) належать також **однорідність** та **регулярність**.

Оцінки бувають точкові і інтервальні. **Точкова оцінка** дає наближене значення невідомого параметру. **Інтервальна оцінка** вказує інтервал, на якому з певною імовірністю знаходиться невідоме значення параметру.

Припустимо, що функція розподілу, яка відповідає вибірці $\xi_1, \xi_2, \dots, \xi_n$, залежить від невідомого параметру θ .

$$P(\xi_k < x) = F_\xi(x, \theta).$$

Оцінкою (точковою оцінкою) $\hat{\theta}_n$ параметру θ називається довільна функція $\hat{\theta}_n = \hat{\theta}_n(\xi_1, \xi_2, \dots, \xi_n)$. Таким чином, $\hat{\theta}_n$ є випадковою величиною. Природньо вимагати, щоб значення оцінки в більшості дослідів були близькі до значення параметру, який оцінюється. Тому говорять про якість оцінювання параметрів.

Для характеристики якості оцінок використовуються такі поняття:

1) **незміщеність**: якщо математичне сподівання оцінки параметру дорівнює самому параметру, тобто

$$M \left\{ \hat{\theta}(\bar{y}) \right\} = \theta ,$$

де \bar{y} — вектор спостережень, $\bar{y} = (y_1, y_2, \dots, y_n)$, то така оцінка називається **незміщеною**.

Приклад 1. Оцінка математичного сподівання —

$$\hat{m}_x = \frac{1}{n} \sum_{i=1}^n x_i .$$

Тоді,

$$M \left\{ \hat{m}_x \right\} = M \left\{ \frac{1}{n} \sum_i x_i \right\} = \int \frac{1}{n} \sum_i x_i p(x) dx .$$

Математичне сподівання (параметр, який оцінюється):

$$m_x = \int x p(x) dx .$$

Утворимо різницю:

$$M\left\{\hat{m}_x\right\} - m_x = \int \left[\frac{1}{n} \sum_i x_i - x\right] p(x) dx .$$

Для того, щоб оцінка була незміщеною, необхідно, щоб ця різниця дорівнювала нулеві. Це буде тоді, коли

$$\frac{\sum_i x_i - nx}{n} = 0 ,$$

тобто, коли $\sum_i x_i = nx$.

Якщо ж математичне сподівання оцінки параметру не дорівнює самому параметру, маємо **зміщену оцінку**.

2) **слушність**: якщо для будь-якого як завгодно малого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left\{\left|\hat{\theta}(y_1, y_2, \dots, y_n)\right| \geq \varepsilon\right\} = 0 ,$$

то така оцінка — **слушна**.

Достатня умова слушності незміщеної оцінки: її дисперсія $\sigma^2 \rightarrow 0$, коли $n \rightarrow \infty$.

3) **ефективність**: якщо $\hat{\theta}_e$ **ефективна оцінка**, то для неї виконується нерівність:

$$M\left\{(\hat{\theta}_e - \theta)^2\right\} \leq M\left\{(\hat{\theta} - \theta)^2\right\} ,$$

де $\hat{\theta}$ — будь-яка інша оцінка.

Для випадку, розглянутому в прикладі 1, $M\left\{(\hat{\theta} - \theta)^2\right\} = \sigma^2$. Дисперсія, що відповідає ефективній оцінці,

$$D_e = M\left\{(\hat{\theta}_e - \theta)^2\right\} = \min\left[M\left\{(\hat{\theta} - \theta)^2\right\}\right] ,$$

звідки можна знайти вигляд ефективної оцінки.

2.6. Сім'я випадкових величин

Нехай на імовірнісному просторі (Ω, \mathcal{F}, P) задані випадкові величини:

$$\xi_1 = \xi_1(\omega), \xi_2 = \xi_2(\omega), \dots, \xi_n = \xi_n(\omega), \omega \in \Omega,$$

які складають *сім'ю випадкових величин*. Кожному ω тут поставлено у відповідність n — вимірний вектор.

Означення 1: *Сумісною функцією розподілу* (або *багатовимірною функцією розподілу*) величин $\xi_1, \xi_2, \dots, \xi_n$ (або випадкового вектора $\xi = (\xi_1, \xi_2, \dots, \xi_n)$) називається імовірність

$$F_{\xi}(x) = F_{\xi_1 \dots \xi_n}(x_1, x_2) = P(\xi_1 < x_1, \xi_2 < x_2),$$

яку розглядають як функцію точки $x = (x_1, \dots, x_n)$ n -вимірного евклідового простору R^n . Функція розподілу $F_{\xi}(x)$ однозначно визначає імовірності $P\{\xi \in B\}$ для довільних паралелепіпедів $B \subset R^n$. Найпростішою з багатовимірних функцій розподілу є двовимірна функція розподілу

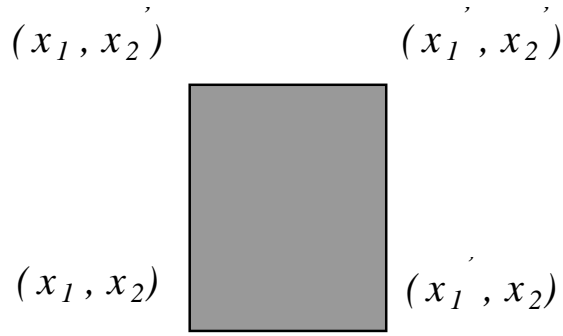
$$F_{\xi}(x) = F_{\xi_1 \xi_2}(x_1, x_2) = P(\xi_1 < x_1, \xi_2 < x_2).$$

На її основі розглянемо властивості багатовимірних функцій розподілу:

1) $0 \leq F(x_1, x_2) \leq 1$;

2) якщо $x_1' > x_1$ і $x_2' > x_2$, то $F(x_1', x_2') \geq F(x_1, x_2)$

3) імовірність для паралелепіпеда (мал. 2.3)



Мал. 2.3.

$$P(x_1 \leq \xi_1 < x_1', x_2 \leq \xi_2 < x_2') = F(x_1', x_2') - F(x_1, x_2') - F(x_1', x_2) + F(x_1, x_2);$$

$$4) F(+\infty, x_2) = F(x_2); F(x_1, +\infty) = F(x_1);$$

$$5) F(x_1, -\infty) = 0; F(-\infty, x_2) = 0; F(-\infty, -\infty) = 0; F(+\infty, +\infty) = 1.$$

Аналогічно одновимірному випадку, якщо випадковий вектор неперервний, можна розглянути таку невід'ємну функцію $p_{\xi}(x) = p_{\xi_1 \dots \xi_n}(x_1, \dots, x_n)$, яка називається **сумісною (або n -вимірною густиною розподілу)**, що для довільного паралелепіпеда B

$$P\{\xi \in B\} = \int\limits_{\substack{K \\ E}} p_{\xi_{IK} \xi_n}(x_1, K, x_n) dx_{IK} dx_n.$$

Розглянемо основні властивості багатовимірної густини розподілу на основі двовимірної:

$$1) p(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2};$$

$$2) p(x_1, x_2) \geq 0;$$

$$3) F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} p(x_1, x_2) dx_1 dx_2 ;$$

$$4) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, x_2) dx_1 dx_2 = 1 ;$$

$$5) \int_{-\infty}^{\infty} p(x_1, x_2) dx_1 = p(x_2) ; \int_{-\infty}^{\infty} p(x_1, x_2) dx_2 = p(x_1) .$$

У ряді випадків використовується умовна функція розподілу

$$F(x_1|x_2) = \frac{F(x_1, x_2)}{F(x_2)}$$

і умовна густина розподілу імовірностей

$$p(x_1|x_2) = \frac{p(x_1, x_2)}{p(x_2)} .$$

Числові характеристики сім'ї випадкових величин.

Розрізняють такі числові характеристики сім'ї випадкових величин (на прикладі двовимірної випадкової величини):

1) початковий момент порядку k :

$$\begin{cases} m_{\xi_1}^k = \int_{-\infty}^{\infty} x_1^k p(x_1) dx_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^k p(x_1, x_2) dx_1 dx_2 ; \\ m_{\xi_2}^k = \int_{-\infty}^{\infty} x_2^k p(x_2) dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2^k p(x_1, x_2) dx_1 dx_2 ; \end{cases}$$

2) центральний момент порядку k :

$$\begin{cases} \mu_{\xi_1}^k = \int_{-\infty}^{\infty} (x_1 - m_{\xi_1})^k p(x_1) dx_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_{\xi_1})^k p(x_1, x_2) dx_1 dx_2; \\ \mu_{\xi_2}^k = \int_{-\infty}^{\infty} (x_2 - m_{\xi_2})^k p(x_2) dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_2 - m_{\xi_2})^k p(x_1, x_2) dx_1 dx_2; \end{cases}$$

3) змішаний початковий момент порядку $k+s$:

$$m_{\xi_1 \xi_2}^{k,s} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^k x_2^s p(x_1, x_2) dx_1 dx_2;$$

4) змішаний центральний момент порядку $k+s$:

$$\mu_{\xi_1 \xi_2}^{k,s} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_{\xi_1})^k (x_2 - m_{\xi_2})^s p(x_1, x_2) dx_1 dx_2.$$

На практиці ж використовують лише такі моменти як:

1) математичне сподівання — початковий момент 1-го порядку:

$$\begin{cases} M_{\xi_1} = m_{\xi_1} = \int_{-\infty}^{\infty} x_1 p(x_1) dx_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 p(x_1, x_2) dx_1 dx_2; \\ M_{\xi_2} = m_{\xi_2} = \int_{-\infty}^{\infty} x_2 p(x_2) dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 p(x_1, x_2) dx_1 dx_2; \end{cases}$$

2) дисперсію — центральний момент 2-го порядку:

$$\begin{cases} D\xi_1 = \mu_{\xi_1}^2 = \int_{-\infty}^{\infty} (x_1 - m_{\xi_1})^2 p(x_1) dx_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_{\xi_1})^2 p(x_1, x_2) dx_1 dx_2; \\ D\xi_2 = \mu_{\xi_2}^2 = \int_{-\infty}^{\infty} (x_2 - m_{\xi_2})^2 p(x_2) dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_2 - m_{\xi_2})^2 p(x_1, x_2) dx_1 dx_2; \end{cases}$$

3) кореляцію — змішаний початковий момент 2-го порядку:

$$r_{\xi_1 \xi_2} = m_{\xi_1 \xi_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2) dx_1 dx_2;$$

4) коваріацію — змішаний центральний момент 2-го порядку:

$$k_{\xi_1 \xi_2} = \text{cov}(\xi_1, \xi_2) = \mu_{\xi_1 \xi_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_{\xi_1})(x_2 - m_{\xi_2}) p(x_1, x_2) dx_1 dx_2;$$

5) коефіцієнт кореляції:

$$\rho(\xi_1, \xi_2) = \frac{\text{cov}(\xi_1, \xi_2)}{\sqrt{D\xi_1 D\xi_2}}.$$

2.7. Евклідов простір випадкових величин

Розглянемо простір елементарних подій $\Omega = \{\omega_i\}$, $i = 1, n$. Якщо задано систему випадкових величин:

$$\xi(\omega_1), \xi(\omega_2), \dots, \xi(\omega_n),$$

то, як вже згадувалось, її зручно представити у вигляді n -вимірного вектора на деякому лінійному просторі, тобто

$$\xi = \xi(\omega) = (\xi(\omega_1), \xi(\omega_2), \dots, \xi(\omega_n)).$$

Для такого простору вводять такі поняття:

— *скалярного добутку* двох векторів —

$$(\xi, \eta) = \sum_{\omega} \xi(\omega) \eta(\omega) p(\omega) = M \xi \eta;$$

тобто, він рівний математичному сподіванню добутку цих векторів;

— *норми* вектора —

$$\|\xi\| = (\xi, \xi)^{\frac{1}{2}} = (M \xi^2)^{\frac{1}{2}};$$

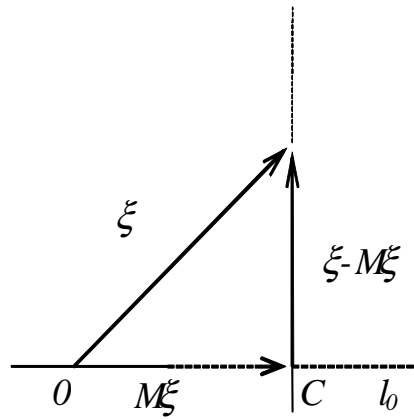
— *метрики* (віддалі між двома векторами) —

$$d(\xi, \eta) = \|\xi - \eta\| = (M(\xi - \eta)^2)^{\frac{1}{2}}.$$

Множина випадкових величин, означених на (Ω, \mathcal{F}, P) , для якої описаним вище способом введені поняття скалярного добутку, норми, метрики, називається багатовимірним **евклідовим простором випадкових величин**. Нехай в n -вимірному евклідовому просторі випадкових величин задано випадковий вектор

$$\xi(\omega) = (\xi(\omega_1), \xi(\omega_2), \dots, \xi(\omega_n))$$

і пряму l_0 (мал. 2.4) $l_0 = \{\xi: \xi(\omega_1) = \xi(\omega_2) = \dots = \xi(\omega_n)\}.$



Мал.2.4.

На прямій l_0 розглянемо деякий вектор m_ξ , $m_\xi \in l_0$, такий, що для довільної точки C , $C \in l_0$,

$$d(\xi, m_\xi) = \min_{C \in l_0} d(\xi, OC).$$

Так як

$$M(\xi - OC)^2 = M(\xi - M\xi)^2 + (M\xi - OC)^2 \geq D\xi,$$

то значення віддалі $d(\xi, OC)$ буде мінімальним при $m_\xi \rightarrow OC = 0$, і тоді $m = M\xi$. Віддаль при цьому:

$$d(\xi, m_\xi) = \sqrt{D}.$$

Отже, мінімальною відстанню між випадковим вектором ξ і прямою l_0 буде відстань між ξ і його математичним сподіванням $M\xi$ на цій прямій, тобто

$$d(\xi, M\xi) = \sqrt{M(\xi - M\xi)^2} = \sqrt{D\xi}.$$

Скалярний добуток:

$$(\xi - M\xi, M\xi) = M\{(\xi - M\xi)M\xi\} = M(\xi M\xi - M^2\xi) = M^2\xi - M^2\xi = 0$$

отже, $(\xi - M\xi, l_0) = 0$, і $(\xi - M\xi) \perp l_0$.

2.8. Кореляція. Геометрична інтерпретація кореляції

Розглянемо два випадкових вектори ξ і η в евклідовому просторі випадкових величин. Представимо їх у вигляді:

$$\begin{aligned}\xi &= \xi_I + M\xi, \\ \eta &= \eta_I + M\eta.\end{aligned}$$

Тоді,

$$\begin{aligned}\xi_I &= \xi - M\xi, \\ \eta_I &= \eta - M\eta.\end{aligned}$$

Знайдемо косинус кута між векторами ξ_I і η_I :

$$\begin{aligned}\cos\varphi_{\xi_I\eta_I} &= \frac{(\xi_I, \eta_I)}{|\xi_I||\eta_I|} = \frac{M\eta_I\xi_I}{\sqrt{M\xi_I^2}M\eta_I^2} = \frac{M(\xi - M\xi)(\eta - M\eta)}{\sqrt{M(\xi - M\xi)^2}(\eta - M\eta)^2} = \\ &= \frac{M(\xi - M\xi)(\eta - M\eta)}{\sqrt{D\xi D\eta}}.\end{aligned}$$

Цей вираз позначають через $\rho(\xi, \eta)$ і називають **коефіцієнтом кореляції**.

Отже, геометричний зміст коефіцієнту кореляції в тому, що це косинус кута між векторами, якими представлені випадкові величини.

Вираз у чисельнику називають **коваріацією** і позначають —

$$\text{cov}(\xi, \eta) = M(\xi - M\xi)(\eta - M\eta).$$

Тоді коефіцієнт кореляції —

$$\rho(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi D\eta}}.$$

Для довільних випадкових векторів ξ і η —

$$|\rho(\xi, \eta)| \leq 1,$$

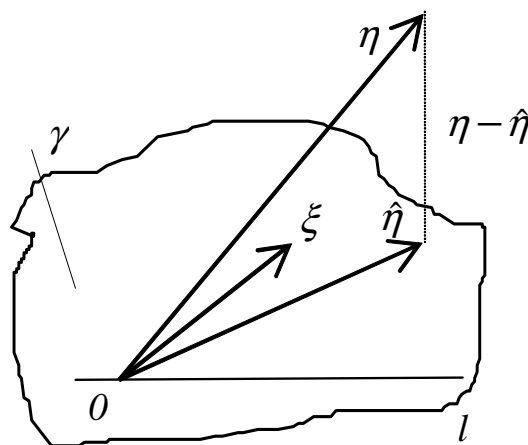
так як, згідно нерівності Коші–Буняковського,

$$(M\xi_l \eta_l)^2 \leq M\xi_l^2 \eta_l^2.$$

Якщо випадкові величини, представлені векторами ξ і η , незалежні, то $\text{cov}(\xi, \eta) = 0$, звідси випливає, що $\rho(\xi, \eta) = 0$ і вектори ξ і η — ортогональні, а відповідні їм випадкові величини — некорельовані. Очевидно, випадкові величини будуть максимально корельованими при $|\rho(\xi, \eta)| = 1$. Причому, при $\rho = 1$ матиме місце коваріантний зв'язок, а при $\rho = -1$ — контрваріантний.

2.9. Рівняння регресії множини випадкових величин

Розглянемо два залежних випадкових вектори ξ і η , причому $\xi \in \gamma$ (мал.2.5).



Мал. 2.5.

Через η позначимо проекцію вектора η на площину γ , $\eta \in \gamma$. Вектор $(\eta - \eta) \perp \gamma$, отже $(\eta - \eta) \perp \xi$, тоді скалярний добуток $(\eta - \eta, \xi) = 0$, тобто

$$M\{(\eta - \eta), \xi\} = 0 \quad (2.9.1)$$

Так як $\eta \in \gamma$ і $\xi \in \gamma$, то можна записати

$$\eta = \alpha\xi + \beta, \quad (2.9.2)$$

де α і β — деякі числа. Рівняння (2.9.2) називають **рівнянням регресії** випадкової величини η на випадкову величину ξ . Знайдемо коефіцієнти α і β . Підставимо (2.9.2) в (2.9.1):

$$\begin{aligned} M\{(\eta - \eta)\xi\} &= M\{(\eta - \alpha\xi - \beta)\xi\} = M\{\eta\xi - \alpha\xi^2 - \beta\xi\} = \\ &= M\eta\xi - \alpha M\xi^2 - \beta M\xi = 0, \end{aligned}$$

звідки

$$M\eta\xi = \alpha M\xi^2 + \beta M\xi \quad (2.9.3)$$

Коефіцієнт кореляції для випадкових величин ξ і η ,

$$\begin{aligned} \rho(\xi, \eta) &= \frac{\text{cov}(\xi, \eta)}{\sqrt{D_\xi D_\eta}} = \frac{M\{(\eta - M\eta)(\xi - M\xi)\}}{\sigma_\eta \sigma_\xi} = \\ &= \frac{M\{\eta\xi - \xi M\eta - \eta M\xi + M\eta M\xi\}}{\sigma_\eta \sigma_\xi} = \frac{M\eta\xi - M\eta M\xi}{\sigma_\eta \sigma_\xi}, \end{aligned}$$

звідки

$$\rho \sigma_\eta \sigma_\xi + M\eta M\xi = M\eta\xi, \quad (2.9.4)$$

де $\sigma_\eta = \sqrt{D\eta}$, $\sigma_\xi = \sqrt{D\xi}$.

Користуючись (2.9.4), можна знайти, що

$$M\xi^2 = \sigma_\xi \sigma_\xi + M\xi M\xi = \sigma_\xi^2 + M^2 \xi. \quad (2.9.5)$$

Підставимо (2.9.5) в (2.9.3) і одержимо —

$$\alpha(\sigma_\xi^2 + M^2 \xi) + \beta M\xi = M\eta\xi$$

або, по-іншому —

$$\alpha\sigma_\xi^2 + (\alpha M\xi + \beta)M\xi = M\eta\xi. \quad (2.9.6)$$

Рівняння (2.9.6) і (2.9.4) об'єднаємо в систему

$$\begin{cases} (\alpha\sigma_\xi) \sigma_\xi + (\alpha M\xi + \beta)M\xi = M\eta\xi; \\ (\rho\sigma_\eta) \sigma_\xi + M\eta M\xi = M\xi\eta. \end{cases} \quad (2.9.7)$$

Очевидно, що $\alpha\sigma_\xi = \rho\sigma_\eta$, $\alpha M\xi + \beta = M\eta$,
звідки

$$\alpha = \rho \frac{\sigma_\eta}{\sigma_\xi}, \quad \beta = M\eta - \alpha M\xi = M\eta - \rho \frac{\sigma_\eta}{\sigma_\xi} M\xi. \quad (2.9.8)$$

Підставляючи (2.9.8) в (2.9.2), одержуємо —

$$\mathfrak{H} = \alpha\xi + \beta = \rho \frac{\sigma_\eta}{\sigma_\xi} \xi + M\eta - \rho \frac{\sigma_\eta}{\sigma_\xi} M\xi = M\eta + \rho \frac{\sigma_\eta}{\sigma_\xi} (\xi - M\xi).$$

Отже, рівняння регресії η на ξ —

$$\mathfrak{H} = M\eta + \rho \frac{\sigma_\eta}{\sigma_\xi} (\xi - M\xi). \quad (2.9.9)$$

За аналогією можна записати рівняння регресії ξ на η :

$$\mathfrak{S} = M\xi + \rho \frac{\sigma_\eta}{\sigma_\xi} (\eta - M\eta). \quad (2.9.10)$$

Знайдемо тепер, за яких умов відстань між векторами η і \mathfrak{H} буде мінімальною, тобто $d(\eta, \mathfrak{H}) = \min$. Використаємо рівняння (2.9.9):

$$\begin{aligned}
d^2(\eta, \xi) &= M(\eta - \xi)^2 = M \left\{ (\eta - M\eta) - \rho \frac{\sigma_\eta}{\sigma_\xi} (\xi - M\xi) \right\}^2 = \\
&= M \left\{ (\eta - M\eta)^2 - 2\rho \frac{\sigma_\eta}{\sigma_\xi} (\xi - M\xi)(\eta - M\eta) + \rho^2 \frac{\sigma_\eta^2}{\sigma_\xi^2} (\xi - M\xi)^2 \right\} = \\
&= D\eta - 2\rho \frac{\sigma_\eta}{\sigma_\xi} \text{cov}(\xi, \eta) + \rho^2 \frac{\sigma_\eta^2}{\sigma_\xi^2} D\xi = D\eta - 2\rho \frac{\sigma_\eta}{\sigma_\xi} \rho \sigma_\xi \sigma_\eta + \rho^2 \sigma_\eta^2 = \\
&= D\eta - 2\rho^2 D\eta + \rho^2 D\eta = D\eta - \rho^2 D\eta = D\eta(1 - \rho^2).
\end{aligned}$$

Як бачимо, $D\eta(1 - \rho^2) = \min = 0$, коли $|\rho| = 1$. В цьому випадку $M(\eta - \xi) = 0$, $\eta = \xi$ і рівняння регресії (2.9.9) набуде вигляду: $\frac{\eta - M\eta}{\sigma_\eta} = \rho \frac{\xi - M\xi}{\sigma_\xi}$.

І на завершення розглянемо дисперсію суми випадкових величин. У випадку двох випадкових величин:

$$\begin{aligned}
D(\xi + \eta) &= M\{\xi + \eta - M(\xi + \eta)\}^2 = M\{(\xi - M\xi) + (\eta - M\eta)\}^2 = \\
&= M\{(\xi - M\xi)^2 + 2(\xi - M\xi)(\eta - M\eta) + (\eta - M\eta)^2\} = \\
&= M(\xi - M\xi)^2 + 2M(\xi - M\xi)(\eta - M\eta) + M(\eta - M\eta)^2 = D\xi + 2\text{cov}(\xi, \eta) + D\eta,
\end{aligned}$$

$$D(\xi + \eta) = D\xi + D\eta + 2\text{cov}(\xi, \eta).$$

Для системи випадкових величин ξ_1, \dots, ξ_n

$$D(\xi_1 + \dots + \xi_n) = \sum_{1 \leq k \leq n} D\xi_k + 2 \sum_{1 \leq k < l \leq n} \text{cov}(\xi_k, \xi_l).$$

2.10. П'яття про збіжність послідовностей випадкових величин

Нехай задано послідовність випадкових величин $\xi_1, \xi_2, \dots, \xi_n, \dots$

Для такої послідовності граничний перехід $\xi \rightarrow \xi_n$ при $n \rightarrow \infty$ можна розуміти по різному, на відміну від граничного переходу для послідовності детермінованих величин, який завжди має один і той самий зміст. Це пояснюється тим, що випадкові послідовності задаються відповідними густинами імовірності.

Отже, для випадкових величин існують різні поняття збіжності. Розглянемо основні з них:

1) збіжність майже напевно:

$$\overset{\text{м.н.}}{\xi_n} \rightarrow \xi,$$

якщо

$$P\left\{\lim_{n \rightarrow \infty} \xi_n = \xi\right\} = 1,$$

або якщо для будь-якого як завгодно малого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left\{\omega: \sup_{m \geq n} |\xi_m - \xi| > \varepsilon\right\} = 0,$$

де *sup* — найменша верхня грань множини;

2) збіжність за імовірністю:

$$\overset{\text{р.}}{\xi_n} \rightarrow \xi,$$

якщо для будь-якого як завгодно малого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\{|\xi_n - \xi| > \varepsilon\} = 0;$$

3) збіжність за розподілом:

—нехай $F_{\xi_n}(x)$ і $F_{\xi}(x)$ — функції розподілу випадкових величин ξ_n і ξ відповідно, тоді:

$$\overset{F.}{\xi_n} \rightarrow \xi,$$

якщо

$$\lim_{n \rightarrow \infty} F_{\xi_n}(x) = F_{\xi}(x).$$

Таку збіжність ще позначають $F_{\xi_n} \rightarrow F_{\xi}$.

Теорема: Якщо $\overset{p.}{\xi_n} \rightarrow \xi$, то $F_{\xi_n}(x) \rightarrow F_{\xi}(x)$, тобто якщо послідовність випадкових величин ξ_n збігається до ξ за імовірністю, то вона збігається до ξ і за розподілом.

4) збіжність в середньому порядку r

$$\overset{r.}{\xi_n} \rightarrow \xi,$$

якщо

$$\lim_{n \rightarrow \infty} M\left\{|\xi_n - \xi|^r\right\} = 0.$$

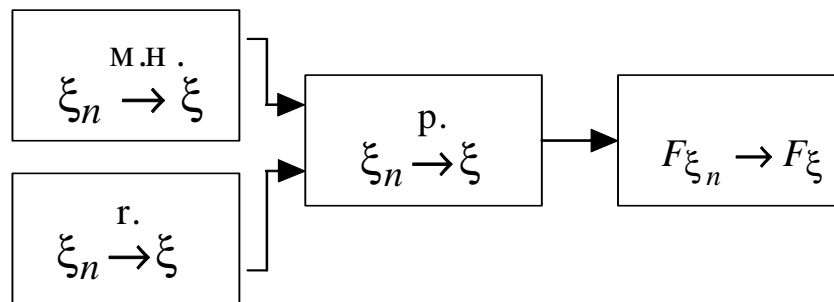
При значенні $r = 1$ маємо просто збіжність в середньому, а при $r = 2$ — збіжність в середньоквадратичному. Якщо $\overset{r=2}{\xi_n} \rightarrow \xi$, то і $\overset{r=1}{\xi_n} \rightarrow \xi$.

Для випадку збіжності в середньому порядку r справедлива нерівність Чебишева

$$P\{|\xi_n - \xi| > \varepsilon\} \geq \frac{M|\xi_n - \xi|^r}{\varepsilon^r}$$

для як завгодно малого $\varepsilon > 0$.

Із всіх розглянутих понять збіжності найчастіше використовується поняття збіжності в середньоквадратичному. Ієрархію різних видів збіжності, яка показано на мал. 2.6.



Мал. 2.6.

Підсумок

Як тільки за елементарну подію обирається числове значення фізичної величини, то спостереження за нею, наприклад, в певний час, дає цілу множину її значень a , отже, і відповідних їм імовірностей. Тому розглядають функцію розподілу імовірностей значень по значеннях величини. Виникає цілий ряд нових понять, пов'язаних з описом таких функцій: математичне сподівання, дисперсія, кореляції (моменти функції розподілу).

Важливим для практики є ще два поняття — збіжності послідовності випадкових величин та якість оцінок моментів за даними експерименту.

Добру наочну інтерпретацію дає “геометризація” понять. Зокрема, приклад цьому — евклідовий простір випадкових величин та рівняння регресії. І помітнішає громіздкість розгляду сім’ї випадкових величин в термінах функцій розподілу.

Розділ 3

ВИПАДКОВІ ПРОЦЕСИ

3.1. Моделі випадкового процесу

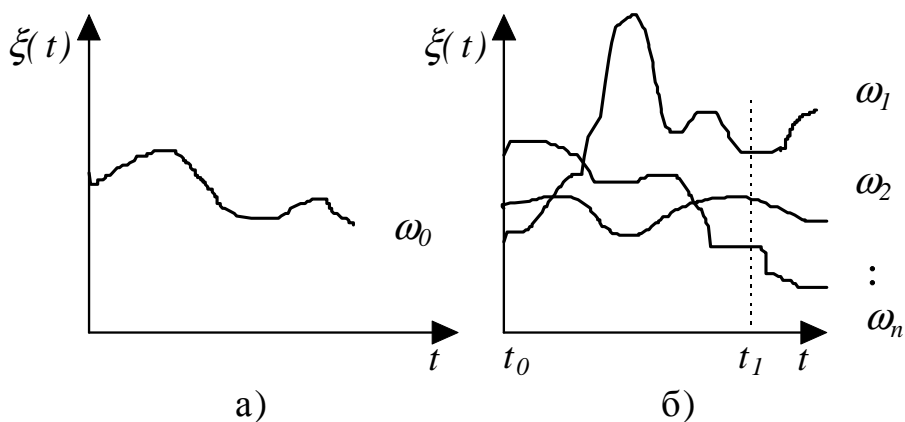
На практиці доводиться часто зустрічатися з параметризованими випадковими величинами, тобто такими, що зв'язані з часом чи просторовими координатами. Прикладами можуть бути власні шуми радіоприймача, зміни тиску, температури, вологості, швидкості вітру та інше. В усіх подібних випадках говорять про параметризований ряд значень випадкової величини, або випадковий (стохастичний, імовірнісний) процес, коли параметром є час, чи поле, коли параметрами є просторові координати.

Випадковим процесом називають сім'ю випадкових величин $\xi(t) = \xi(\omega, t)$, $\omega \in \Omega, t \in \mathbb{R}$, заданих на імовірнісному просторі $(\Omega, \mathcal{F}, \mathbb{P})$. Параметр t інтерпретують як час. Дійсну функцію t при фіксованому ω називають **реалізацією** випадкового процесу (застосовують також терміни "траєкторія" випадкового процесу і "вибіркова функція"). Якщо ж зафіксувати t , то матимемо звичайну випадкову величину.

Можливі два підходи до вивчення випадкових процесів:

1) розглядають лише одну "довгу" реалізацію випадкового процесу, тобто $\xi(t) = \xi(\omega_0, t)$ (мал.3.1.,а);

2) розглядають велику кількість "коротких" (t_0, t_1) реалізацій випадкового процесу (так званий **ансамбль реалізацій**), тобто $\xi(t) = \xi(\omega, t)$ (мал.3.1.,б).



Мал. 3.1.

Якщо результати цих розглядів “однакові”, то такий процес називають *ергодичним*.

Так як випадковий процес $\xi(t)$ при фіксованих значеннях t є випадковою величиною, то для його опису застосовують ті ж імовірнісні характеристики, що і для випадкових величин, а саме: функції розподілу імовірностей, густини імовірностей, характеристичні функції, моменти.

1. Функції розподілу і густини імовірності. Для ансамблю реалізацій випадкового процесу одновимірна функція розподілу імовірності —

$$F_1(x_1, t_1) = P\{\xi(t_1) < x_1\},$$

і одновимірна густина імовірності —

$$p_1(x_1, t_1) = \frac{\partial}{\partial x_1} F_1(x_1, t_1).$$

Слово "одновимірна" підкреслює той факт, що розглядаються значення випадкового процесу в один фіксований момент часу t_1 . Тому одновимірна функція розподілу і густина імовірності не є повними характеристиками випадкового процесу, вони ніби описують процес "статично", в даний момент часу.

Більш повними характеристиками випадкового процесу є двовимірна функція розподілу:

$$F_2(x_1, x_2; t_1, t_2) = P\{\xi(t_1) < x_1, \xi(t_2) < x_2\}$$

і двовимірна густина імовірності:

$$p_2(x_1, x_2; t_1, t_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F_2(x_1, x_2; t_1, t_2).$$

У загальному випадку двовимірна функція розподілу і густина імовірності також не дають вичерпного опису випадкового процесу.

Для більш повного і детального опису необхідно розглядати багатовимірну функцію розподілу:

$$F_n(x_1, \dots, x_n; t_1, \dots, t_n) = P\{\xi(t_1) < x_1, \dots, \xi(t_n) < x_n\}$$

і, відповідно, багатовимірну густину імовірності:

$$p_n(x_1, \dots, x_n; t_1, \dots, t_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F(x_1, \dots, x_n; t_1, \dots, t_n).$$

Отже, випадковий процес у загальному випадку описується за допомогою n -вимірної, або нескінченно вимірної функцій розподілу (густини розподілу).

Для сумісного опису двох чи декількох випадкових процесів використовують сумісні функції розподілу і густини розподілу імовірності. Так, для двох процесів $\xi(t)$ і $\eta(t)$

$$\begin{aligned} F_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) &= \\ &= P(\xi(t_1) < x_1, \dots, \xi(t_n) < x_n, \eta(t'_1) < y_1, \dots, \eta(t'_m) < y_m); \\ p_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) &= \\ &= \frac{\partial^{n+m}}{\partial x_1 \dots \partial x_n \partial y_1 \dots \partial y_m} F_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m). \end{aligned}$$

Якщо виконується умова:

$$\begin{aligned} p_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) &= \\ &= p_n(x_1, \dots, x_n; t_1, \dots, t_n) p_m(y_1, \dots, y_m; t'_1, \dots, t'_m), \end{aligned}$$

то випадкові процеси $\xi(t)$ і $\eta(t)$ називаються **незалежними**.

2. Характеристична функція. Поряд з функціями розподілу і густиною імовірності, для опису випадкового процесу $\xi(t_i)$ використовують характеристичну функцію:

$$\begin{aligned}\varphi_n(z_1, \dots, z_n; t_1, \dots, t_n) &= M\{\exp(iz_1\xi_1 + \dots + iz_n\xi_n)\} = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp[i(z_1x_1 + \dots + z_nx_n)] p_n(x_1, \dots, x_n; t_1, \dots, t_n) dx_1 \dots dx_n,\end{aligned}$$

тут $\xi_i = \xi(t_i)$.

3. Моментні функції. Для опису випадкових процесів використовуються моментні функції. Розрізняють:

1) початкові моментні функції різних порядків:

— k_1 -го порядку:

$$m_{\xi_1}^{k_1}(t) = \int x^{k_1} p_1(x; t) dx ;$$

— $(k_1 + k_2)$ -го порядку:

$$m_{\xi_1, \xi_2}^{k_1, k_2}(t_1, t_2) = \iint x_1^{k_1} x_2^{k_2} p_2(x_1, x_2; t_1, t_2) dx_1 dx_2 ;$$

— $(k_1 + \dots + k_n)$ -го порядку:

$$\begin{aligned}m_{\xi_1, \dots, \xi_n}^{k_1, \dots, k_n}(t_1, \dots, t_n) &= \\ &= \int \dots \int x_1^{k_1} \dots x_n^{k_n} p_n(x_1, \dots, x_n; t_1, \dots, t_n) dx_1 \dots dx_n ;\end{aligned}$$

2) центральні моментні функції різних порядків:

— k_1 -го порядку:

$$\mu_{\xi_1}^{k_1}(t) = \int [x - m_{\xi_1}(t)]^{k_1} p_1(x; t) dx ;$$

— $(k_1 + k_2)$ -го порядку:

$$\begin{aligned} \mu_{\xi_1, \xi_2}^{k_1, k_2}(t_1, t_2) &= \\ &= \int \dots \int [x_1 - m_{\xi_1}(t_1)]^{k_1} [x_2 - m_{\xi_2}(t_2)]^{k_2} p_2(x_1, x_2; t_1, t_2) dx_1 dx_2 ; \\ &\text{— } (k_1 + \dots + k_n)\text{-го порядку:} \end{aligned}$$

$$\begin{aligned} \mu_{\xi_1, \dots, \xi_n}^{k_1, \dots, k_n}(t_1, \dots, t_n) &= \\ &= \int \dots \int [x_1 - m_{\xi_1}(t_1)]^{k_1} \dots [x_n - m_{\xi_n}(t_n)]^{k_n} p_n(x_1, \dots, x_n; t_1, \dots, t_n) dx_1 \dots dx_n . \end{aligned}$$

На практиці використовують також змішані моменти:

1) **коваріаційну функцію** — початковий момент 2-го порядку

$$K_{\xi}(t_1, t_2) = m_{\xi}^{1,1}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p_2(x_1, x_2; t_1, t_2) dx_1 dx_2 ;$$

$$K_{\xi}(t_1, t_2) = M\xi(t_1)\xi(t_2) ;$$

2) **кореляційну функцію** — центральний момент 2-го порядку

$$\begin{aligned} R_{\xi}(t_1, t_2) &= m_{\xi}^{1,1}(t_1, t_2) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - m_{\xi}(t_1)][x_2 - m_{\xi}(t_2)] p_2(x_1, x_2; t_1, t_2) dx_1 dx_2 ; \end{aligned}$$

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

Зауважимо, що між моментними функціями випадкових величин (випадкових процесів) і моментами для сімей випадкових величин існує термінологічна невідповідність: коваріаційна функція процесу за

своїм виглядом відповідає кореляції (коваріаційному моменту), а кореляційна функція — коваріації (кореляційному моменту) випадкових величин.

Між коваріаційною і кореляційною функціями випадкового процесу існує зв'язок:

$$K_{\xi}(t_1, t_2) = R_{\xi}(t_1, t_2) + m_{\xi}(t_1)m_{\xi}(t_2).$$

Розділ теорії, який вивчає випадкові процеси для означення яких достатньо тільки властивостей, які визначаються лише цими характеристиками, називається *кореляційною теорією*.

3.2. Неперервність та диференційовність випадкових процесів

Як уже згадувалось, випадковий процес $\xi(t)$ — це сім'я випадкових величин, що залежать від параметру t ($t \in [a, b]$, $t \in (a, b)$ або $t \in (-\infty, \infty)$), який інтерпретується як час. Розглянемо випадкові процеси, які приймають, взагалі кажучи, комплексні значення і для яких $M|\xi(t)|^2 < \infty$. Значення процесу в момент часу t будемо інтерпретувати як елемент простору $L_2 = L_2(\Omega, F, P)$. Тоді процес $\xi(t)$ можна розглядати як деяку криву в просторі L_2 .

Нехай задано сім'ю випадкових величин $\{\xi_{\lambda}, \lambda \in \Lambda\}$, де $\xi_{\lambda} \in L_2$, а $\xi_{\lambda} \in \xi_{\lambda}(t)$.

Означення 1: Функція

$$K(\lambda, \lambda') = M \xi_{\lambda} \overline{\xi_{\lambda'}} = (\xi_{\lambda}, \xi_{\lambda'})$$

називається *коваріаційною функцією* сімей (ξ_{λ}) .

Властивості:

$$1) K(\lambda, \lambda') = M |\xi_{\lambda}|^2 \geq 0;$$

$$2) K(\lambda, \lambda') = K(\overline{\lambda'}, \lambda);$$

$$3) |K(\lambda, \lambda')|^2 \leq K(\lambda, \lambda)K(\lambda', \lambda') \text{ — нерівність Коші-Буняковського};$$

$$4) \sum_{k,r=1}^n c_k \overline{c_r} k(\lambda_k, \lambda_r) \geq 0 \text{ де } c_k, c_r \text{ — комплексні коефіцієнти.}$$

Означення 2: Функція

$$R(\lambda, \lambda') = M(\xi_\lambda - m(\lambda))(\overline{\xi_{\lambda'} - m(\lambda')}) ,$$

де $m(\lambda) = M\xi_\lambda$, $m(\lambda') = M\xi_{\lambda'}$, називається **кореляційною функцією** сімей (ξ_λ) випадкових величин.

Наведемо означення і умови неперервності та диференційовності випадкового процесу.

Нехай на множині Λ означено додатню функцію $g(\lambda)$, $g(\lambda) > 0$, таку, що її найбільша нижня границя рівна нулеві:

$$\inf(g(\lambda), \lambda \in \Lambda) = 0.$$

Означення 3: Границею в середньоквадратичному сімей випадкових величин $\{\xi_\lambda, \lambda \in \Lambda\}$, $\xi_\lambda \in L_2$ при $g(\lambda) \rightarrow 0$ називають випадкову функцію η , $\eta \in L_2$, таку, що для будь-якого $\varepsilon > 0$ знайдеться таке $\delta = \delta(\varepsilon) > 0$, що $M|\eta - \xi_\lambda|^2 < \varepsilon$ для всіх λ , для яких $g(\lambda) < \delta$.

Якщо $\eta \in L_2$ — середньоквадратична границя сімей (ξ_λ) , то пишуть $\xi_\lambda \xrightarrow{L_2} \eta$ при $g(\lambda) \rightarrow 0$, або

$$\eta = \lim_{g(\lambda) \rightarrow 0} \xi_\lambda .$$

Теорема 1: Для існування середньоквадратичної границі η сім'ї (ξ_λ) необхідно і достатньо, щоб існувала границя:

$$\lim_{g(\lambda)+g(\lambda') \rightarrow 0} K(\lambda, \lambda') = K_0 .$$

Якщо ця границя існує, то $K_0 = M|\eta|^2$.

Означення 4: Випадковий процес $\{\xi(t), t \in (a, b)\}$, $\xi(t) \in L_2$ називається:

а) **неперервним в середньоквадратичному** в точці t , якщо:

$$\lim_{|t'-t|\rightarrow 0} \xi(t') = \xi(t) ;$$

б) *диференційовним в середньоквадратичному*, якщо існує границя:

$$\lim_{|h|\rightarrow 0} \frac{\xi(t+h) - \xi(t)}{h} = \xi'(t) .$$

Випадкову величину $\xi'(t)$, якщо вона існує, називають *середньоквадратичною похідною* процесу $\xi(t)$ в точці t . Якщо процес неперервний в середньоквадратичному в кожній точці інтервалу (a,b) , то його називають неперервним в середньоквадратичному на (a,b) . Аналогічно, якщо процес диференційовний в середньоквадратичному в кожній точці інтервалу (a,b) , то він диференційовний в середньоквадратичному на (a,b) .

Теорема 2:

а) Для неперервності в середньоквадратичному процесу $\xi(t)$ в точці t необхідно і достатньо, щоб:

$$K(t',t') - K(t',t) - K(t,t') + K(t,t) \rightarrow 0$$

при $t' \rightarrow t$.

б) Для диференційовності в середньоквадратичному процесу $\xi(t)$ в точці t необхідно і достатньо, щоб існувала границя —

$$\lim_{t' \rightarrow t, t'' \rightarrow t} \frac{K(t',t'') - K(t',t) - K(t,t'') + K(t,t)}{(t'-t)(t''-t)} = \frac{\partial^2 K(t,s)}{\partial t \partial s} \Big|_{s=t} .$$

У загальному випадку така границя, якщо вона існує, називається

узагальненою похідною $\frac{\partial^2 K(t,s)}{\partial t \partial s} \Big|_{s=t} .$

3.3. Інтегрування випадкових процесів

Розглянемо випадковий процес $\xi(t)$, $\xi(t) \in L_2$, на деякому відрізку $[a, b]$. Для цього розіб'ємо відрізок $[a, b]$ на менші відрізки за допомогою точок: t_1, \dots, t_n , $a = t_0 < t_1 < \dots < t_n = b$, $\Delta t_k = t_{k+1} - t_k$. Позначимо $|\lambda| = \max |\Delta t_k|$.

Означення 1: Інтегралом

$$\int_a^b \xi(t) dt$$

від випадкового процесу $\xi(t)$, $t \in [a, b]$, $\xi(t) \in L_2$, називають границю інтегральних сум $\lim_{|\lambda| \rightarrow 0} l.i.m. S_\lambda$, де:

$$S_\lambda = \sum_k \xi(t_k) \Delta t_k .$$

Для існування інтегралу достатньо, щоб коваріаційна функція $K(t, s)$ була інтегровна за Ріманом на квадраті $t \geq a$, $s \leq b$. При цьому:

$$M \left| \int_a^b \xi(t) dt \right|^2 = \int_a^b \int_a^b K(t, s) dt ds .$$

Якщо випадковий процес $\xi(t)$ заданий на нескінченному відрізку $[a, \infty)$, то вводять поняття невластного інтегралу —

$$\int_a^\infty \xi(t) dt = \lim_{b \rightarrow \infty} \int_a^b \xi(t) dt .$$

Для його існування достатньо, щоб функція $K(t, s)$ була інтегровна за Ріманом на довільному квадраті $[a, b] \times [a, b]$, $b > a$ і існував невластний інтеграл:

$$\int_a^\infty \int_a^\infty K(t, s) dt ds .$$

Аналогічно можна ввести поняття інтегралу від $-\infty$ до $+\infty$.
 Дають означення інтегралу вигляду:

$$I = \int_a^b f(t) d\xi(t) ,$$

де $\xi(t)$ — випадковий процес, причому дане означення повинно бути аналогічним означенню звичайного інтегралу Стільт'єса. В такому випадку на процес $\xi(t)$ накладають певні умови — зокрема, він повинен бути процесом з ортогональними приростами.

Випадковий процес $\xi(t)$, $t \in (a, b)$, називають **процесом з ортогональними приростами**, якщо $\xi(t) \in L_2$ і для довільних t_1, t_2, t_3, t_4 ($a < t_1 < t_2 < t_3 < t_4 < b$)

$$M[\xi(t_4) - \xi(t_3)][\overline{\xi(t_2) - \xi(t_1)}] = 0 .$$

У випадку, коли $f(t)$ — не випадкова функція, а $\xi(t)$ — випадковий процес з ортогональними приростами, випадкову величину $I(t)$ називають **стохастичним інтегралом**

$$I(t) = \int_{-\infty}^t f(t) d\xi(t) .$$

3.4. Стаціонарний випадковий процес. Спектральна густина потужності стаціонарного процесу

Важливий вид випадкових процесів складають **стаціонарні процеси**. Так називають процеси, імовірнісні характеристики яких не змінюються з часом. Розрізняють стаціонарність у вузькому і широкому сенсі.

Означення 1: Випадковий процес $\xi(t)$, $t \in (-\infty, \infty)$, називається **стаціонарним у вузькому сенсі**, якщо всі багатовимірні функції

розподілу не залежать від часу, тобто при довільних n і t_0 справедлива рівність:

$$F_n(x_1, \dots, x_n; t_1 + t_0, \dots, t_n + t_0) = F_n(x_1, \dots, x_n; t_1, \dots, t_n).$$

Стаціонарний у вузькому сенсі випадковий процес називають ще **сильно стаціонарним** або **стаціонарним процесом першого порядку**.

Означення 2: Випадковий процес $\xi(t)$, $t \in (-\infty, \infty)$, називається **стаціонарним у широкому сенсі**, якщо при $M|\xi(t)|^2 < \infty$ його математичне сподівання не залежить від часу (є постійною величиною), а кореляційна функція залежить тільки від різниці аргументів $(t_2 - t_1)$, тобто:

$$M\xi(t) = m = \text{const},$$

$$R(t_1, t_2) = M(\xi(t_1) - m)(\overline{\xi(t_2) - m}) = R(t_2 - t_1).$$

Стаціонарний у широкому сенсі випадковий процес називають ще **слабо стаціонарним**, або **стаціонарним процесом другого порядку**.

Випадковий процес, стаціонарний у вузькому сенсі, завжди стаціонарний і в широкому сенсі. Обернене твердження в загальному випадку невірне.

Якщо випадковий процес $\xi(t)$ розглядати як вектор, то із означення слабкої стаціонарності випливає, що:

1) норма вектора $\xi(t)$ не залежить від часу:

$$|\xi(t)| = \sqrt{(\xi(t), \xi(t))} = \sqrt{M|\xi(t)|^2} = \sqrt{K(0)};$$

2) кут φ між векторами $\xi(t_1)$ і $\xi(t_2)$, $t_2 = t_1 + \tau$, є неперервною функцією від τ , $\tau = t_2 - t_1$,

$$\cos \varphi = \frac{(\xi(t_1), \xi(t_1 + \tau))}{|\xi(t_1)| |\xi(t_1 + \tau)|} = \frac{K(\tau)}{K(0)},$$

де $K(\tau) = K(t_1, t_1 + \tau) = M\xi(t_1)\overline{\xi(t_1 + \tau)}$ — коваріаційна функція.
Розглянемо процес вигляду:

$$\xi(t) = \sum_{k=1}^n \gamma_k e^{iu_k t}. \quad (3.4.1)$$

Він є сумою простих гармонічних коливань

$$\xi_k(t) = \gamma_k e^{iu_k t} = a_k e^{i\varphi_k} e^{iu_k t} = a_k e^{i(\varphi_k + u_k t)}.$$

Тут a_k — амплітуда, $\gamma_k = a_k e^{i\varphi_k}$ — комплексна амплітуда, φ_k — початкова фаза, u_k — частота гармонічного коливання $\xi_k(t)$. Сукупність амплітуд на частотах u_k , $k=1, \dots, n$, які розглядають як множину точок на прямій $(-\infty, \infty)$, називають **спектром** процесу $\xi(t)$. Будемо вважати комплексні амплітуди γ_k випадковими величинами. Тоді $\xi(t)$ — випадковий комплекснозначний процес, $t \in (-\infty, \infty)$.

Припустимо, що комплексні амплітуди γ_k мають нульові середні значення і між собою не корельовані, тобто

$$M\gamma_k = 0; M\gamma_k \bar{\gamma}_r = 0, \text{ коли } k \neq r; \quad (3.4.2)$$

$$M\gamma_k \bar{\gamma}_r = M|\gamma_k|^2 = C_k^2, \text{ коли } k = r.$$

Тоді:

$$\begin{aligned} M\xi(t) &= \sum_{k=1}^n M\xi_k(t) = 0; \\ R(t_1, t_2) &= M(\xi(t_1) - M\xi(t_1))(\overline{\xi(t_2) - M\xi(t_2)}) = M\xi(t_1)\overline{\xi(t_2)} = \\ &= M\left[\sum_{k,r=1}^n \xi_k(t_1)\overline{\xi_r(t_2)}\right] = M\left[\sum_{k,r=1}^n \gamma_k e^{iu_k t_1} \bar{\gamma}_r e^{-iu_r t_2}\right] = \\ &= M\left[\sum_{k,r=1}^n \gamma_k \bar{\gamma}_r e^{i(u_k t_1 - u_r t_2)}\right] = \sum_{k=1}^n C_k^2 e^{iu_k(t_1 - t_2)}. \end{aligned}$$

Таким чином, кореляційна функція процесу $\xi(t)$ залежить тільки від різниці $t_1 - t_2$, $R(t_1, t_2) = R(t_1 - t_2) = R(\tau)$, і задається формулою:

$$R(t) = \sum_{k=1}^n C_k^2 e^{iuk t}. \quad (3.4.3)$$

Отже, якщо виконані умови (3.4.2), то випадковий процес $\xi(t)$ є стаціонарним у широкому сенсі. Формулу (3.4.3) називають **спектральним зображенням** (представленням) кореляційної функції.

Функцію

$$F(u) = \sum_{uk < u} C_k^2,$$

називають **спектральною функцією** процесу (3.4.1). Вона є сумою величин C_k^2 гармонічних складових процесу $\xi(t)$, частоти яких менші заданого значення u . За допомогою спектральної функції кореляційну функцію процесу $\xi(t)$ можна записати у вигляді:

$$R(t) = \int_{-\infty}^{\infty} e^{iut} dF(u). \quad (3.4.4)$$

Виявляється, що формула (3.4.4) може бути поширена на випадкові процеси й іншого вигляду.

Теорема 1 (Хінчина): Для того, щоб функція $R(t)$ була кореляційною функцією неперервного в середньоквадратичному слабо стаціонарного процесу, необхідно і достатньо, щоб вона допускала зображення (3.4.4).

Означення 3: Функцію $F(u)$, яка міститься у формулі (3.4.4), називають **спектральною функцією слабо стаціонарного процесу**. Якщо $F(u)$ абсолютно неперервна, тобто:

$$F(u) = \int_{-\infty}^u f(v) dv,$$

то функцію $f(u)$ називають **спектральною густиною** процесу.

Якщо кореляційна функція $R(t)$ абсолютно інтегровна на інтервалі $(-\infty, \infty)$

$$\int_{-\infty}^{\infty} |R(t)| dt < \infty,$$

то можна записати, що:

$$f(u) = \int_{-\infty}^{\infty} e^{iut} R(t) dt,$$

тобто, спектральна густина рівна перетворенню Фур'є від кореляційної функції.

3.5. Приклади стаціонарних випадкових процесів

Розглянемо деякі приклади випадкових процесів і покажемо, що вони стаціонарні.

1. Гармонічне коливання з випадковими параметрами. Нехай заданий випадковий процес:

$$\xi(t) = A(t) \cos(\omega t + \varphi),$$

де φ — випадкова початкова фаза рівномірно розподілена на інтервалі $(-\pi, \pi)$, $A(t)$ — випадкова амплітуда, яка не залежить від φ і є стаціонарним у широкому сенсі випадковим процесом, тобто

$$\begin{aligned} MA(t) &= m_A = \text{const}; \\ R_A(t_1, t_2) &= R_A(t_2 - t_1) = R_A(\tau). \end{aligned} \quad (3.5.1, 3.5.2)$$

Знайдемо математичне сподівання і кореляційну функцію для $\xi(t)$. Так як випадковий процес $A(t)$ та φ незалежні, то

$$\begin{aligned} M\xi(t) &= MA(t)M\{\cos(\omega t + \varphi)\} = MA(t)M\{\cos\omega t \cos\varphi - \sin\omega t \sin\varphi\} = \\ &= MA(t)[M\{\cos\varphi\}\cos\omega t - M\{\sin\varphi\}\sin\omega t]. \end{aligned}$$

Тут

$$M\{\sin \varphi\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin \varphi d\varphi = \frac{1}{2\pi} \cos \varphi \Big|_{-\pi}^{\pi} = 0,$$

$$M\{\cos \varphi\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos \varphi d\varphi = \frac{1}{2\pi} \sin \varphi \Big|_{-\pi}^{\pi} = 0;$$

тому

$$M\xi(t) = 0 = \text{const} = m_{\xi}. \quad (3.5.3)$$

Кореляційна функція

$$\begin{aligned} R_{\xi}(t_1, t_2) &= M(\xi(t_1) - m_{\xi})(\xi(t_2) - m_{\xi}) = M\xi(t_1)\xi(t_2) = \\ &= M\{A(t_1)A(t_2)\cos(\omega t_1 + \varphi)\cos(\omega t_2 + \varphi)\} = \\ &= MA(t_1)A(t_2)M\{\cos(\omega t_1 + \varphi)\cos(\omega t_2 + \varphi)\} = \\ &= MA(t_1)A(t_2)M\left\{\frac{1}{2} [\cos(\omega t_1 - t_2) + \cos(\omega t_1 + \omega t_2 + 2\varphi)]\right\} = \\ &= \frac{1}{2} MA(t_1)A(t_2) [\cos(\omega t_1 - t_2) + M\{\cos(\omega t_1 + \omega t_2 + 2\varphi)\}]. \end{aligned}$$

У даному виразі $MA(t_1)A(t_2) = K_A(t_1, t_2)$ — коваріаційна функція процесу $A(t)$. З (3.5.1) і (3.5.2) випливає, що:

$$K_A(t_1, t_2) = R_A(t_1, t_2) + m_A(t_1)m_A(t_2) = R_A(\tau) + m_A^2 = K_A(\tau).$$

Математичне сподівання

$$\begin{aligned} M\{\cos(\omega t_1 + \omega t_2 + 2\varphi)\} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(\omega t_1 + \omega t_2 + 2\varphi) d\varphi = \\ &= \frac{1}{2} \frac{1}{2\pi} \sin(\omega t_1 + \omega t_2 + 2\varphi) \Big|_{-\pi}^{\pi} = 0. \end{aligned}$$

Тому —

$$\begin{aligned} R_{\xi}(t_1, t_2) &= \frac{1}{2} K_A(\tau) \cos \omega(t_1 - t_2) = \frac{1}{2} K_A(\tau) \cos \omega(t_1 - t_2) = \\ &= \frac{1}{2} K_A(\tau) \cos \omega\tau = R_{\xi}(\tau). \end{aligned} \quad (3.5.4)$$

Вирази (3.5.3) і (3.5.4) свідчать про те, що $\xi(t)$ — стаціонарний в широкому сенсі випадковий процес. Отже, при заданих умовах, гармонічне коливання із випадковими параметрами є стаціонарним в широкому сенсі випадковим процесом.

2. Гармонічне коливання з випадковими параметрами в загальному випадку. Запишемо гармонічне коливання в загальному вигляді:

$$\xi(t) = A(t)e^{i(\omega t + \varphi)}. \quad (3.5.5)$$

Тут випадкові величини $A(t)$ і φ відповідають тим самим параметрам, що і в попередньому прикладі.

Математичне сподівання випадкового процесу $\xi(t)$:

$$\begin{aligned} M\xi(t) &= MA(t)Me^{i(\omega t + \varphi)}; \\ Me^{i(\omega t + \varphi)} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(\omega t + \varphi)} d\varphi = \frac{1}{2\pi} e^{i\omega t} \int_{-\infty}^{\infty} e^{i\varphi} d\varphi = \\ &= \frac{1}{2\pi} e^{i\omega t} \frac{1}{i} e^{i\varphi} \Big|_{-\pi}^{\pi} = \frac{1}{2\pi i} e^{i\omega t} (\cos \varphi + i \sin \varphi) \Big|_{-\pi}^{\pi} = 0; \end{aligned}$$

$$M\xi(t) = 0 = \text{const} = m_{\xi}.$$

Кореляційна функція у нашому випадку:

$$\begin{aligned} R_{\xi}(t_1, t_2) &= M(\xi(t_1) - m_{\xi})(\overline{\xi(t_2) - m_{\xi}}) = M\xi(t_1)\overline{\xi(t_2)} = \\ &= M\{A(t_1)A(t_2)e^{i(\omega t_1 + \varphi)}e^{-i(\omega t_2 + \varphi)}\} = M\{A(t_1)A(t_2)e^{-i\omega(t_2 - t_1)}\} = \\ &= e^{-i\omega(t_2 - t_1)}MA(t_1)MA(t_2) = e^{-i\omega\tau}K_A(\tau) = R_{\xi}(\tau). \end{aligned}$$

Отже, гармонічне коливання, задане виразом (3.5.5), також є стаціонарним в широкому сенсі випадковим процесом.

3. Суперпозиція випадкових гармонічних коливань.

Гармонічне коливання, розглянуте в попередньому прикладі, можна записати і так:

$$\xi(t) = A(t)e^{i(\omega t + \varphi)} = A(t)e^{i\varphi}e^{i\omega t} = \gamma e^{i\omega t},$$

де $\gamma = A(t)e^{i\varphi}$ — комплексна амплітуда, яка являє собою випадкову величину; ω — частота гармонічного коливання.

Утворимо суперпозицію таких коливань з різними частотами:

$$\xi(t) = \sum_{k=1}^n \xi_k(t) = \sum_{k=1}^n \gamma_k e^{i\omega_k t}. \quad (3.5.6)$$

Припустимо, що комплексні амплітуди γ_k мають нульові середні значення і між собою не корельовані, тобто:

$$M\gamma_k = 0; M\gamma_k \bar{\gamma}_r = 0, \text{ коли } k \neq r; \quad (3.5.7)$$

$$M\gamma_k \bar{\gamma}_r = M|\gamma_k|^2 = C_k^2, \text{ коли } k = r.$$

Тоді

$$\begin{aligned} M\xi(t) &= \sum_{k=1}^n M\xi_k(t) = 0 = \text{const} = m_\xi; \\ R_\xi(t_1, t_2) &= M(\xi(t_1) - m_\xi)(\overline{\xi(t_2) - m_\xi}) = M\xi(t_1)\overline{\xi(t_2)} = \\ &= M\left\{ \sum_{k,r=1}^n \xi_k(t_1)\overline{\xi_r(t_2)} \right\} = M\left\{ \sum_{k,r=1}^n \gamma_k e^{i\omega_k t_1} \overline{\gamma_r e^{i\omega_r t_2}} \right\} = \\ &= M\left\{ \sum_{k,r=1}^n \gamma_k \bar{\gamma}_r e^{-i(\omega_r t_2 - \omega_k t_1)} \right\} = \sum_{k=1}^n C_k^2 e^{-i\omega_k(t_2 - t_1)} = \\ &= \sum_{k=1}^n C_k^2 e^{-i\omega_k \tau} = R_\xi(\tau). \end{aligned}$$

Отже, суперпозиція випадкових гармонічних коливань, для яких виконуються умови (3.5.7), являє собою стаціонарний в широкому сенсі випадковий процес.

4. "Білий шум". Досить часто при розгляді випадкових процесів використовують поняття "білого шуму".

Білим шумом називають узагальнений випадковий процес $\xi(t)$, $t \in (-\infty, \infty)$, з некорельованими значеннями і постійною спектральною густиною

$$f(u) = \sigma^2, \quad u \in (-\infty, \infty).$$

Так як такий процес має рівномірний спектр в дуже широкому діапазоні частот, його і прийнято називати "білим шумом", по аналогії з "білим світлом", яке має у видимій частині рівномірний суцільний спектр.

Кореляційна функція білого шуму —

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(u) e^{iu\tau} du = \frac{\sigma^2}{2\pi} \int_{-\infty}^{\infty} e^{iu\tau} du.$$

У теорії функцій доводиться, що інтеграл в правій частині $\int_{-\infty}^{\infty} e^{iu\tau} du = 2\pi\delta(t)$, тому

$$R(\tau) = \frac{\sigma^2}{2\pi} 2\pi\delta(t) = \sigma^2(t).$$

Отже, кореляційна функція білого шуму має вигляд δ -імпульсу. Виходячи з цього, можна сказати, що **білий шум** — це випадковий процес з некорельованими значеннями або δ -корельований процес.

Білий шум не існує реально, так як в природі не існує сигналів із рівномірним спектром у діапазоні частот $(-\infty, \infty)$. При обчисленні середньої потужності білого шуму одержуємо нескінченність, що фізично неправдоподібно. Білий шум слід розглядати тільки як

математичну узагальнену функцію, яка при відповідних умовах використовується для опису реальних шумів.

Білий шум — це стаціонарний процес. Він може бути стаціонарним як у широкому, так і вузькому сенсі. Якщо для випадкового процесу $\xi(t)$, $t \in (-\infty, \infty)$, значення некорельовані, тобто $R(\tau) = \delta(\tau)$, то такий процес називають білим шумом, стаціонарним в широкому сенсі, або просто **білим шумом в широкому сенсі**. Якщо ж значення випадкового процесу незалежні (не існують вищі моменти, ніж другий), то матимемо **білий шум у вузькому сенсі** (стаціонарний у вузькому сенсі).

3.6. Лінійні перетворення стаціонарних випадкових процесів

Будемо розглядати деяку лінійну ланку, на вхід якої поступає функція $f(t)$, а з виходу знімається функція $g(t)$. Співвідношення між $f(t)$ і $g(t)$ можна записати у вигляді:

$$g(t) = Tf(t),$$

де T — деякий оператор.

Система називається лінійною, якщо:

- 1) клас функцій, які допустимі на вході системи, лінійний;
- 2) оператор T відповідає принципу суперпозиції

$$T(a_1 f_1 + a_2 f_2) = a_1 T f_1 + a_2 T f_2,$$

де a_1 і a_2 — константи.

Введемо оператор зсуву в часі:

$$\theta_a f(t) = f(t + a), a \in (-\infty, \infty).$$

Якщо для системи виконуються умови:

$$\theta_a f(t) = f(t) \text{ і } T\theta_a = \theta_a T,$$

де $f(t)$ — будь-яка із допустимих на вході системи функція, то така система називається **інваріантною в часі**.

Найчастіше зустрічаються лінійні перетворення вигляду:

$$g(t) = \int_{-\infty}^{\infty} h(t, s) f(s) ds. \quad (3.6.1)$$

Функцію $h(t, s)$ в цьому виразі називають **імпульсною ваговою функцією**.

Для інваріантних в часі систем імпульсна вагова функція:

$$h(t, s) = h(t + a, s + a) = h(t - s, 0) = h(t - s),$$

і вираз (3.6.1) набуває вигляду:

$$g(t) = \int_{-\infty}^{\infty} h(t - s) f(s) ds. \quad (3.6.2)$$

Інтеграл такого вигляду називається згорткою і записується:

$$g = h * f.$$

У реальних фізичних системах, якщо вхідна функція $f(t) = 0$ при $t \leq t_0$, то і функція на виході $g(t) = 0$ при $t \leq t_0$. Тому імпульсна вагова функція реальної системи повинна відповідає **умові фізичної реалізованості**

$$h(t, s) = 0 \text{ при } t < s.$$

Для інваріантної системи ця умова набуває вигляду:

$$h(t) = 0 \text{ при } t < 0.$$

Надалі будемо розглядати лінійні інваріантні системи. Як відомо, для таких систем вираз (3.6.2) у частотній області має вигляд:

$$G(iu) = H(iu)F(iu),$$

де $G(iu)$ і $F(iu)$ — перетворення Фур'є від функцій $g(t)$ і $f(t)$ відповідно; $H(iu)$ — перетворення Фур'є від функції $h(t)$, або функція передачі (інакше: **частотна характеристика** або **коефіцієнт передачі**).

У деяких випадках замість перетворення Фур'є зручно використовувати перетворення Лапласа, тобто

$$G(p) = H(p)F(p),$$

$$\text{де } H(p) = \int_{-\infty}^{\infty} h(t)e^{-pt} dt, p = ju, j = \sqrt{-1}.$$

Нехай тепер на вхід інваріантної лінійної системи з імпульсною ваговою функцією $h(t)$ подається слабо стаціонарний процес $\xi(t)$, $M\xi(t) = 0$. Процес на виході системи —

$$\eta(t) = \int_{-\infty}^{\infty} h(t-s)\xi(s)ds. \quad (3.6.3)$$

Для існування інтегралу достатньо, щоб:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(s, s') ds ds' < \infty,$$

де $K(s, s')$ — коваріаційна функція процесу $h(t-s)\xi(s)$,

$$\begin{aligned} K(s, s') &= Mh(t-s)\xi(s)\overline{h(t-s')\xi(s')} = h(t-s)\overline{h(t-s')}M\xi(s)\overline{\xi(s')} = \\ &= h(t-s)\overline{h(t-s')}K_{\xi}(s, s') = h(t-s)\overline{h(t-s')}R_{\xi}(s-s'), \end{aligned}$$

де $R_{\xi}(s-s') = R_{\xi}(\tau)$ — кореляційна функція процесу $\xi(t)$.

Скористаємось спектральним зображенням кореляційної функції —

$$R_{\xi}(t) = \int_{-\infty}^{\infty} e^{iut} dF(u) = \int_{-\infty}^{\infty} e^{iut} F(du),$$

одержимо:

$$\begin{aligned}
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(s, s') ds ds' &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t-s) \overline{h(t-s')} \int_{-\infty}^{\infty} e^{iu(s-s')} F(du) ds ds' = \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t-s) e^{-iu(t-s)} \int_{-\infty}^{\infty} \overline{h(t-s')} e^{iu(t-s')} ds' ds F(du) = \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t-s) e^{-iu(t-s)} \overline{H(iu)} ds F(du) = \int_{-\infty}^{\infty} H(iu) \overline{H(iu)} F(du) = \\
&= \int_{-\infty}^{\infty} |H(iu)|^2 F(du).
\end{aligned}$$

Отже, якщо:

$$\int_{-\infty}^{\infty} |H(iu)|^2 F(du) < \infty, \quad (3.6.4)$$

то інтеграл (3.6.3) існує для всіх $t > 0$. При цьому $M\eta(t)=0$, а кореляційна функція —

$$\begin{aligned}
R_{\eta}(t, t') &= M(\eta(t) - M\eta(t))(\overline{\eta(t') - M\eta(t')}) = M\eta(t)\overline{\eta(t')} = \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t-s, t'-s') ds ds' = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t-s) \overline{h(t'-s')} R_{\xi}(s-s') ds ds' = \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t-s) \overline{h(t'-s')} \int_{-\infty}^{\infty} e^{iu(s-s')} F(du) ds ds' = \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t-s) e^{-iu(t-s)} \int_{-\infty}^{\infty} \overline{h(t'-s')} e^{iu(t'-s')} ds' ds e^{iu(t-t')} F(du) = \\
&= \int_{-\infty}^{\infty} e^{iu(t-t')} |H(iu)|^2 F(du).
\end{aligned}$$

Таким чином, на виході системи одержуємо також слабо стаціонарний процес з кореляційною функцією

$$R_{\eta}(t) = \int_{-\infty}^{\infty} e^{iut} |H(iu)|^2 F(du), \quad (3.6.5)$$

де $F(u)$ — спектральна функція процесу на вході системи.

Означення 1: Перетворення T називають **фільтром**, якщо воно має вигляд (3.6.3) і виконана умова (3.6.4), або якщо воно є середньоквадратичною границею таких перетворень.

Теорема 1: Для того, щоб функція $H(iu)$ була частотною характеристикою фільтру для процесу $\xi(t)$ із спектральною функцією $F(u)$, необхідно і достатньо, щоб вона задовільняла умову (3.6.4). Кореляційна функція процесу на виході фільтру з частотною характеристикою $H(iu)$ задається формулою (3.6.5).

Теорема 2: Якщо частотні характеристики фільтрів $H_1(iu)$ і $H_2(iu)$ відповідають умові (3.6.4), $\eta_1(t)$, $\eta_2(t)$ — процеси на виході відповідних фільтрів, то у випадку фільтру з частотною характеристикою —

$$H_3(iu) = H_1(iu) + H_2(iu),$$

процес на його виході —

$$\eta_3(t) = \eta_1(t) + \eta_2(t),$$

то процеси $\eta_1(t)$, $\eta_2(t)$ є стаціонарно зв'язані, з взаємно кореляційною функцією

$$R_{12}(t) = \int_{-\infty}^{\infty} e^{iut} H_1(iu) \overline{H_2(iu)} F(du).$$

3.7. Спектральний розклад слабо стаціонарного процесу

Спектральний розклад кореляційної функції

$$R(t) = \int_{-\infty}^{\infty} e^{iut} F(du),$$

де $F(u) = \sum_{u_k < u} C_k^2 = \sum_{u_k < u} M|\gamma_k|^2$, настановує на думку, що існує можливість

і слабо стаціонарний процес $\xi(t)$ розкласти по гармоніках γe^{iut} , $u \in (-\infty, \infty)$. Тоді $F(du)$ можна інтерпретувати як енергію, що переноситься гармонічними коливаннями з частотами в інтервалі $(u, u + du)$, а формула

$$R_\eta(t) = \int_{-\infty}^{\infty} e^{iut} |H(iu)|^2 F(du),$$

показує, що при проходженні випадкового процесу $\xi(t)$ через фільтр енергія його гармонічних складових з частотами в інтервалі $(u, u + du)$ домножується на $|H(iu)|^2$. Про це свідчать наступні теореми.

Теорема 1: Нехай $\xi(t)$ — слабо стаціонарний процес із спектральною функцією $F(u)$, $M\xi(t) = 0$. Тоді існує процес $\zeta(u)$, $u \in (-\infty, \infty)$, з ортогональними приростами такий, що $M|\xi(u)|^2 = F(u)$ і

$$\xi(t) = \int_{-\infty}^{\infty} e^{iut} \zeta(du). \quad (3.7.1)$$

Формулу (3.7.1) називають **спектральним розкладом** слабо стаціонарного процесу $\xi(t)$, а функцію $\zeta(\cdot)$ — **стохастичною спектральною мірою** цього процесу.

Теорема 2: Нехай слабо стаціонарний процес $\xi(t)$ із спектральним розкладом (3.7.1) подається на вхід фільтру із частотною характеристикою $H(iu)$. Тоді процес $\eta(t)$ на виході фільтру має спектральний розклад:

$$\eta(t) = \int_{-\infty}^{\infty} e^{iut} H(iu) \zeta(du). \quad (3.7.2)$$

3.8. Стаціонарні випадкові послідовності

Коли параметр t стаціонарного випадкового процесу є дискретним параметром, то говорять про стаціонарні випадкові послідовності. Розглянемо деякі з таких послідовностей.

1. Нехай маємо випадковий процес $\xi(t)$ з функцією розподілу $F_{\xi}(t)$ такий, що:

$$M\xi(t) = m_{\xi} = 0, \sigma_{\xi}(t) = \sigma_{\xi} = 1, t = 0, \pm 1, \pm 2, \dots$$

і коваріаційна функція:

$$K(s) = M\xi(t+s)\xi(t) = \begin{cases} 1, & s = 0; \\ 0, & s \neq 0. \end{cases}$$

Такий випадковий процес називають **послідовністю некорельованих випадкових величин**.

2. Нехай випадковий процес задається формулою:

$$x(t) = \sum_{k=0}^n a_k \xi(t-k) = a_0 \xi(t) + a_1 \xi(t-1) + \dots + a_n \xi(t-n), t = 0, \pm 1, \pm 2, \dots,$$

де $a_k, k=0, \dots, n$, — деякі числа.

Якщо $Mx(t) = m_x = 0, \sigma_x(t) = \sigma_x = 1$, то коваріаційна функція

$$K(s) = \begin{cases} \sum_{k=0}^{n-|s|} a_{k+|s|} a_k, & |s| \leq n; \\ 0, & |s| > n, \end{cases}$$

і такий процес називається **послідовністю ковзного середнього порядку n** .

3. Узагальнимо попередній випадок. Покладемо $n \rightarrow \infty$. Тоді, якщо виконується умова

$$\sum_{k=0}^{\infty} |a_k|^2 < \infty,$$

то випадковий процес

$$x(t) = \sum_{k=0}^n a_k \xi(t-k) = a_0 \xi(t) + a_1 \xi(t-1) + \dots + a_n \xi(t-n), t = 0, \pm 1, \pm 2, \dots,$$

з коваріаційною функцією

$$K(s) = Mx(t+s)x(t) = \sum_{k=0}^{\infty} a_{k+s} a_k$$

називається просто **послідовністю ковзного середнього**.

4. Розглянемо випадковий процес вигляду:

$$x(t) = ax(t-1) + c\xi(t), t = 0, \pm 1, \pm 2, \dots,$$

де a і c — дійсні числа;

$$\sigma^2 = Dc\xi(t) = |c|^2.$$

Такий процес називається **процесом авторегресії 1-го порядку**.

Можна розглянути і процес **авторегресії m -го порядку**:

$$x(t) = \bar{a}_1 x(t-1) + \bar{a}_2 x(t-2) + \dots + \bar{a}_m x(t-m) + c\xi(t),$$

або, по-іншому —

$$x(t) + a_1 x(t-1) + a_2 x(t-2) + \dots + a_m x(t-m) = -c\xi(t),$$

Для коваріаційної функції в цьому випадку можна записати:

$$K(s) + a_1 K(s-1) + a_2 K(s-2) + \dots + a_m K(s-m) = |c|^2.$$

Процес авторегресії ще називають процесом Юла-Уокера.

5. Якщо випадковий процес містить в собі процеси ковзного середнього і авторегресії, тобто має вигляд:

$$\begin{aligned} x(t) + a_1 x(t-1) + a_2 x(t-2) + \dots + a_m x(t-m) = \\ = -[c\xi(t) + b_1 \xi(t-1) + \dots + b_n \xi(t-n)], t = 0, \pm 1, \pm 2, \dots, \end{aligned}$$

то такий процес називається **змішаним процесом авторегресії-ковзного середнього**.

Підсумок

Розглянуто поняття, які мають аналогіями поняття матаналізу — аналізу детермінованих функцій. Важливо, що в означення неперервності, а отже, диференційовності, інтегровності і, очевидно, складніших операторів входять елементи опису випадкових функцій — розподілу, густини розподілу та їх параметрів. Показано, що можливо (у випадку стаціонарності, ергодичності) розвинути методи представлення випадкових процесів, подібні до спектральних розкладів періодичних і інтегровних з квадратом детермінованих функцій. Лінійні перетворення випадкових процесів за виглядом їх математичного опису нагадують лінійні перетворення детермінованих процесів, проте, використовують для цього характеристики випадкових процесів — кореляційні функції, спектральні густини потужностей та інші. Зображення самих процесів пов'язані з поняттям випадкової міри. Якщо процеси з незалежними (ортогональними) приростами, то легко одержати і їх лінійні перетворення, проте, все одно практично застосовними є кореляційні чи спектральні характеристики результату перетворень.

У цьому розділі розглянуто лише питання прикладного застосування випадкових процесів у рамках теорії другого порядку (спектрально-кореляційної). У деяких випадках розвинуто питання прикладного застосування теорії випадкових процесів з використанням функції розподілу, коли останні мають певні властивості, до яких відносять марківську властивість — залежність стохастичних властивостей відліку лише від стохастичних властивостей попереднього відліку, а також їх граничні випадки (марківські неперервні процеси).

Розділ 4

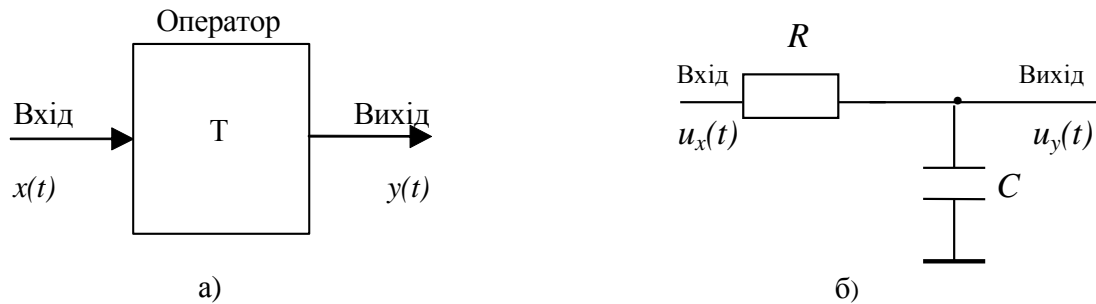
МАТЕМАТИЧНІ ОСНОВИ ЛІНІЙНОЇ ОБРОБКИ СИГНАЛІВ В АПАРАТУРІ РАДІОЕЛЕКТРОННИХ БІОМЕДИЧНИХ СИСТЕМ

4.1. Структура математичної моделі апаратури для обробки біосигналів

Біооб'єкту властиві зміни у часі різних біофізичних величин (температури, тиску, концентрації розчину тощо). Якщо в характеристиках або параметрах цих величин є відомості про стан біооб'єкту, то називатимемо їх біосигналами. Оскільки на конкретну біофізичну величину впливають різні чинники, то при біомедичних дослідженнях виникає необхідність в обробці (фільтрації) відповідного біосигналу для зменшення результату впливу небажаних чинників, виділення бажаних характеристик біосигналу тощо. Для побудови біомедичних систем обробки біосигналів використовують математичні моделі біосигналів (адекватні їм математичні об'єкти) і, відповідно, математичні моделі радіоелектронної апаратури. Існують різні методології математичного моделювання. За однією з них обґрунтовується заміна біотехнічної системи (або її ланки) математичним об'єктом — лінійним оператором (фільтром, мал.4.1). При цьому математичний об'єкт для представлення оператора набирає вигляду залежно від еквівалентних (у математичному сенсі) представлень біосигналу — у власному та ізоморфному йому енергетичному просторах, або у просторі змінних стану (табл. 4.1). Вигляд представлення біосигналів вибирається з практичних міркувань. Від нього залежать складність теоретичної (синтезу схем) та технічної (конструювання апаратів) побудови ланок системи, її метрологічні характеристики (точність реалізації характеристик її функцій та чутливість їх до малих змін їх параметрів), надійність. Математичною моделлю ланки системи є операторне рівняння, наприклад:

$$y(t) = T\{x(t)\}, \quad (4.1.1)$$

де x — функція дійсної змінної (математична модель вхідного біосигналу), y — функція дійсної змінної (математична модель обробленого біосигналу), T — оператор (математична модель обробки біосигналів). Для багатьох входів та виходів вираз (4.1.1) набирає векторно-матричного вигляду.



Мал. 4.1. (а) — функціональна схема реалізації оператора T ; (б) — принципова схема реалізації оператора в апаратурі (RC- ланкою, u — напруга)

Таблиця 4.1

Математичні моделі

Простір математичного моде- лювання біосигналів	Математичні моделі	
	Біосигналів	Ланок систем
Власний	Неперервні функції дійсної змінної (часу), часові послідовності	Оператор: а) диференціальний (диференціальне рів- няння), різницевий (різницеве рівняння); б) інтегральний (згортка з ядром дифе- ренціального опера- тора)
Енергетичний	Функція комплексної змінної (частоти)	Оператор — дробово- лінійна комплексно значна функція
Фазовий (змінних станів)	Вектори неперервних функцій — входу,	Оператор — система лінійних диферен-

	виходу, змінних станів тощо	ціальних векторно- матричних рівнянь
--	--------------------------------	---

Система (або її ланка) називається **стаціонарною**, якщо її параметри та функція не залежать від часу. Стаціонарні ланки та системи називають системами з постійними в часі параметрами. Якщо ж властивості системи залежать від часу, то вона не буде стаціонарною (тоді це система із змінними в часі параметрами, або параметрична система).

Якщо для ланки (системи) справедливі принципи суперпозиції та однорідності:

$$y = T\{ax_1(t) + ax_2(t)\} = \underbrace{T\{ax_1(t)\} + T\{ax_2(t)\}}_{\text{суперпозиція}} = \underbrace{a(T\{x_1(t)\} + T\{x_2(t)\})}_{\text{однорідність}}, \quad (4.1.2)$$

де a — константа, то така ланка (система) називається **лінійною**. Інакше така ланка (система) нелінійна. Система що складається з лінійних ланок є лінійною. Але, лінійна система не завжди складається з лінійних ланок (тобто, можлива лінеаризація системи). Подібне твердження справедливе для стаціонарності.

Якщо математична модель (4.1.1) біосигналу x та апаратури його обробки T , разом з властивостями (4.1.2)) забезпечує формулювання і вирішення завдання побудови цієї апаратури, то ця математична модель є **адекватною** до біосигналу.

Для обробки біосигналу насамперед виконується його відбір від біооб'єкта та перетворення в електричну величину, передача цієї величини до пристрою обробки. Цей факт, як правило, спеціально не наголошується у літературі з обробки біосигналів. У подальшому електронною апаратурою виконується: згладжування (низькочастотна фільтрація), виділення (селекція) сигналу з шумів і завад (субоптимальна або оптимальна фільтрація), інтерполяція та екстраполяції (прогноз його поведінки) тощо. У всіх цих випадках вибір, синтез характеристик обробки та розрахунків їх параметрів (параметричний синтез) виконуються з застосуванням відповідного **критерію**. Критерієм є функціонал від характеристики обробки, який є математично коректним та фізично інтерпретованим, і будується (або вибирається) за біофізичною інтерпретацією математичного змісту

задачі синтезу, такий, що набирає відповідного екстремального значення (максимуму або мінімуму, залежно від змісту задачі) тоді, коли результат обробки задовільний. Наприклад, для лінійного оператора обробки коректним математично (як метрика лінійного функційного простору) та інтерпретованим фізично (як віддаль) таким функціоналом є середнє значення квадрату похибки.

Приклади:

а) Фільтрація (селекція) сигналу з адитивної суміші сигналу та шуму, імовірнісні характеристики яких відомі, *оптимальна* за критерієм мінімальної середньоквадратичної похибки (фільтр Колмогорова-Вінера). Вона виконується лінійною ланкою з функцією передачі (імпульсною ваговою функцією), визначеною за допомогою згаданого критерію шляхом розв'язування відповідної оптимізаційної задачі;

б) Оптимальне оцінювання (також за квадратичним критерієм) неспостережних станів біооб'єкту за відомими спостережними станами при заданих імовірнісних характеристиках сигналів та шумів (шляхом обчислювальної рекурсивної процедури — фільтр Калмана);

Обробкою сигналів виконується їх спектральний аналіз, визначаються їх статистики — функції розподілу ймовірностей та їх моменти (математичне сподівання, дисперсія), кореляційні функції тощо.

У біомедичній практиці застосування результатів обробки біосигналів розбиваються на два класи: а) „ручні” (коли застосування знаходять результати обробки що є оцінками морфологічних параметрів біосигналу) й інтерактивні (коли розв'язують завдання у діалозі з біотехнічною системою за такими параметрами); б) автоматизовані (завдання розв'язує система).

У даному розділі наведено приклади характерних методів обробки випадкових сигналів у лінійних стаціонарних біотехнічних системах.

4.2. Вплив лінійної ланки на функцію розподілу ймовірностей значень біосигналу

Біосигнал на вході ланки системи може бути заданим його стохастичними характеристиками. Тоді при обробці сигналів у лінійних системах виникають задачі:

а) визначення за відомими багатовимірними функціями розподілу біосигналу на вході системи функцій розподілу сигналу на виході;

б) визначення густини спектру потужності, автокореляційної функції вихідного сигналу за відомими такими характеристиками вхідного біосигналу.

Для розв'язку задач типу (б) використовують методи спектрально-кореляційної теорії сигналів та їх перетворень. Розв'язок задачі типу (а) досить громіздкий. Відомі окремі випадки:

1) якщо випадковий процес на вході лінійної системи має нормальний закон розподілу, то процес на виході також буде нормальним;

2) якщо на вході системи процес, відмінний від нормального, то і на виході одержимо не нормальний процес, який, крім того, відрізнятиметься від вхідного. Практично важливим є випадок, коли незалежно від густини розподілу імовірності процесу на вході системи можна стверджувати, що вихідний процес має нормальний розподіл. Це явище називається **нормалізацією** випадкових процесів і характерне для лінійних інерційних систем.

Нормалізація випадкових процесів має місце тоді, коли ширина енергетичного спектру вхідного сигналу Δf_{ξ} значно більша, ніж смуга пропускання лінійної системи Δf , тобто:

$$\Delta f_{\xi} \gg \Delta f ,$$

або

$$\frac{\tau}{\tau_0} = a \gg 1,$$

де τ — постійна часу системи, τ_0 — інтервал кореляції,

$$\tau_0 = \frac{1}{R(0)} \int_0^{\infty} R(\tau) d\tau,$$

де $R(\tau)$ — кореляційна функція.

Чим більше a , тим більше закон розподілу вихідного сигналу наближається до нормального.

Явище нормалізації є прямим наслідком центральної граничної теореми Ляпунова: якщо x_1, x_2, \dots, x_n — незалежні випадкові величини, які мають скінченні Mx_i і Dx_i , то закон розподілу суми —

$$Y = \sum_{i=1}^n x_i$$

при збільшенні n наближається до нормального.

Випадковий сигнал на виході лінійної системи

$$\eta(t) = \int_0^{\infty} \xi(\tau) h(t - \tau) d\tau,$$

$$\eta(t) = \sum_k \xi(k\Delta\tau) h(t - k\Delta\tau) \Delta\tau.$$

Інтервал кореляції τ_0 у випадку широкосмугового процесу $\xi(t)$ значно менший, ніж $\Delta\tau$, так що будь-які дві випадкові величини $\xi(k\Delta\tau)$ і $\xi((k+1)\Delta\tau)$ — практично незалежні. Тому процес на виході системи, який рівний сумі великої кількості незалежних з однаковою інтенсивністю випадкових величин, по теоремі Ляпунова матиме розподіл, що наближається до нормального.

4.3. Кореляційна функція і енергетичний спектр відгуку лінійної ланки системи на стаціонарний випадковий процес

Позначимо сигнал на вході системи через $x(t)$, а сигнал на її виході — $y(t)$. Перетворення Фур'є від цих сигналів —

$$X(u) = \int_{-\infty}^{\infty} x(t) e^{-iut} dt,$$

$$Y(u) = \int_{-\infty}^{\infty} y(t) e^{-iut} dt. \quad (4.3.1)$$

Величини $X(u)$ і $Y(u)$ зв'язані між собою залежністю

$$Y(u) = H(iu)X(iu), \quad (4.3.2)$$

де $H(iu)$ — функція передачі даної системи. Виконавши обернене перетворення Фур'є від (4.3.2), одержимо значення вихідного сигналу в момент часу t :

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(iu)X(iu)e^{iut} du. \quad (4.3.3)$$

Множник $\frac{1}{2\pi}$ у виразі (4.3.3) іноді опускають. Очевидно, значення вихідного сигналу в момент часу $t + \tau$ дорівнюватиме:

$$y(t + \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(iu')X(iu')e^{iu't}e^{iu'\tau} du'. \quad (4.3.4)$$

Функція $y(t)$ дійсна, тому формула (4.3.4) не зміниться, якщо в її правій частині перейти до комплексно-спряжених величин:

$$y(t + \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{H(iu')} \overline{X(iu')} e^{-iu't} e^{-iu'\tau} du'. \quad (4.3.5)$$

Процес $x(t)$ — стаціонарний, отже $y(t)$ також буде стаціонарним, і $m_x = m_y = 0$. Кореляційна функція вихідного сигналу в цьому випадку:

$$\begin{aligned} R_y(\tau) = My(t)y(t + \tau) &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(iu) \overline{H(iu')} \times \\ &\times M\{X(iu) \overline{X(iu')}\} e^{it(u-u')} e^{-iu'\tau} du du'. \end{aligned} \quad (4.3.6)$$

В спектрально-кореляційній теорії доводиться, що для стаціонарного процесу

$$M\{X(iu) \overline{X(iu')}\} = 2\pi F_x(u) \delta(u - u'),$$

де $F_x(u)$ — **спектральна густина** або **енергетичний спектр** сигналу $x(t)$.

Використовуючи фільтруючу властивість δ -функції, формулу (4.3.6) можна спростити —

$$R_y(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_x(u) |H(iu)|^2 e^{-iu\tau} du,$$

або, що те саме,

$$R_y(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_x(u) |H(iu)|^2 e^{iu\tau} du, \quad (4.3.7)$$

Таким чином, ми одержали вираз для кореляційної функції відгуку лінійної системи на стаціонарний випадковий процес.

Згідно теореми Вінера-Хінчина, кореляційна функція стаціонарного процесу може бути знайдена за допомогою оберненого перетворення Фур'є від його спектральної густини. Для сигналу $y(t)$ це запишеться

$$R_y(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_y(u) e^{iu\tau} du. \quad (4.3.8)$$

Порівнюючи формули (4.3.7) і (4.3.8), робимо висновок, що:

$$F_y(u) = F_x(u) |H(iu)|^2,$$

тобто енергетичний спектр вихідного сигналу рівний енергетичному спектрові вхідного сигналу, помноженому на квадрат модуля функції передачі системи.

4.4. Вплив на білий шум інтегруючої ланки

4.4.1. Білий шум та вінерівський процес.

В областях застосування статистичної радіотехніки (управління, зв'язок, контроль та інше) досить часто адекватними моделями збурень є білий шум та вінерівський процес.

Білий шум — узагальнений випадковий процес $\{n(t), t \in T\}$ (noise — шум) з нулевим математичним сподіванням і некорельованими значеннями:

$$M[n(t)] = 0, \quad M[n(t)n(t+\tau)] = \frac{N_0}{2} \delta(\tau),$$

де $\delta(\tau)$ — дельта-функція Дірака, так що кореляційна функція білого шуму є узагальненою функцією. Параметр $N_0 > 0$, визначає інтенсивність білого шуму. У випадку залежності його від часу матимемо нестационарний білий шум. Стационарний білий шум можна представити як:

$$n(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} dz(\omega), \quad (4.4.1.1)$$

де $z(\omega)$ — випадковий процес з некорельованими приростами. Іноді кажуть, що стохастичний процес, який має зображення (4.4.1.1), є гармонізований за Лоевом (бо використано гармонічне коливання $e^{i\omega t}$), але в узагальненому сенсі. Спектральну густину $S(\omega)$ білого шуму знайдено за теоремою Вінера-Хінчина:

$$S(\omega) = \int_{-\infty}^{\infty} \frac{N_0}{2} \delta(\tau) e^{-j\omega\tau} d\tau = \frac{N_0}{2},$$

тобто, вона є постійною величиною, що означає нескінченну енергію. Тому білий шум є математично ідеалізованим, абстрактним процесом. Проте, застосування при аналізові реальних фізичних систем білого шуму дає правильні результати, тому що відгук реальної фізичної системи на δ -подібне збурення обмежений у часі. Цю властивість і відображає назва білого шуму як узагальненого процесу.

Розглянемо процес вигляду:

$$\eta(t) = \int_0^t n(\tau) d\tau, \quad t \in [0, \infty]$$

або:

$$\eta'(t) = n(t).$$

Прирости $\eta(t)$ на неперетинних інтервалах часу є незалежними. Тобто, якщо існує границя $\lim_{|\tau-s| \rightarrow 0} \frac{\eta(\tau) - \eta(s)}{\tau - s}$ (похідна), то вона є значеннями білого шуму $n(t)$. Але така похідна в загальному не існує, хоч у “дограничному” сенсі, розглядаючи незалежні прирости, можна твердити про фізичне існування такого процесу, а розглядаючи неперервний час (здійснюючи граничний перехід), говоримо тоді про узагальнений випадковий процес (білий шум).

Незалежність приростів на неперетинних інтервалах часу означає марковість процесу $\Delta\eta \equiv \eta(\tau) - \eta(s)$, $s - \tau \equiv \Delta t_i$, $\Delta t_i \cap \Delta t_j \in \emptyset$, $i \neq j$ (процес називають марковим (марківським)), якщо його майбутнє значення залежить від теперішнього і не залежить від минулого (в імовірнісному сенсі). Тоді вираз набуває наступного вигляду: $\vartheta_1(\eta_i, t_i | \eta_1, t_1; \eta_{2i}, t_{2i}; \dots; \eta_{i-1}, t_{i-1}) = \vartheta_2(\eta_i, t_i | \eta_{i-1}, t_{i-1})$, ϑ — умовна імовірність значень η , t — час.

Оскільки білий шум у неперервному часі є узагальненим випадковим процесом, то він не описується функцією розподілу, такої функції просто не існує. Проте, розглядаючи процес $\Delta\eta$, можна розглядати функції розподілів імовірності значень процесу η . Якщо цей розподіл гаусовий $w(\eta, \Delta t) = \frac{1}{\sqrt{\pi N_0 \Delta t}} \exp \left\{ -\frac{\eta^2}{N_0 \Delta t} \right\}$, то його математичне сподівання $m_{\Delta\eta}(t)$ рівне нулеві (бо $m_n(t) = 0$, а дисперсія (центральный момент другого порядку)

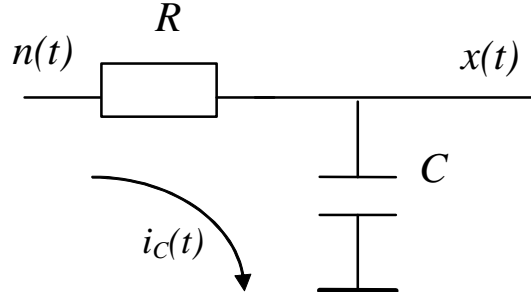
$$\sigma_{\Delta\eta}^2 = m \left[\int_0^t n(\tau) d\tau \int_0^t n(s) ds \right] = \int_0^t \int_0^t m[n(\tau)n(s) d\tau ds] = \frac{N_0}{2} \int_0^t \int_0^t \delta(\tau - s) d\tau ds = \frac{N_0}{2} \Delta t,$$

що означає ріст дисперсії з часом.

Нормальний (гаусовий) марківський процес $\eta(t)$ з стаціонарними і незалежними приростами називають **вінеровим (вінерівським) процесом**. Білий шум не є марковим процесом. Але незалежність його значень при як завгодно малих проміжках часу між ними не суперечить умові марковості.

4.4.2. Визначення основних характеристик процесу на виході RC-ланки

Розглянемо RC-ланку, коли на її вхід подано білий шум (мал.4.2).



Мал. 4.2. $x(t)$ — напруга на виході RC-ланки, $n(t)$ — напруга на вході RC-ланки, значення якої моделюють білим шумом (від джерела напруги, внутрішній опір $r_{вн} = 0$).

Оскільки $i_c(t) = \frac{dx(t)}{dt}$, то рівняння, яке описує процеси в RC-ланці, матиме вигляд: $\frac{dx}{dt} = \alpha x + \alpha \eta$, де $\alpha = \frac{1}{RC}$, аргументи опущено. Нехай початковою умовою є $x(t_0) = x_0$, тоді розв'язок такого рівняння матиме вигляд:

$$x(t) = e^{-\alpha(t-t_0)} \left[x_0 + \alpha \int_{t_0}^t e^{\alpha\tau} n(\tau) d\tau \right].$$

При $t_0 = 0$ середнє значення (математичне сподівання) процесу на виході буде —

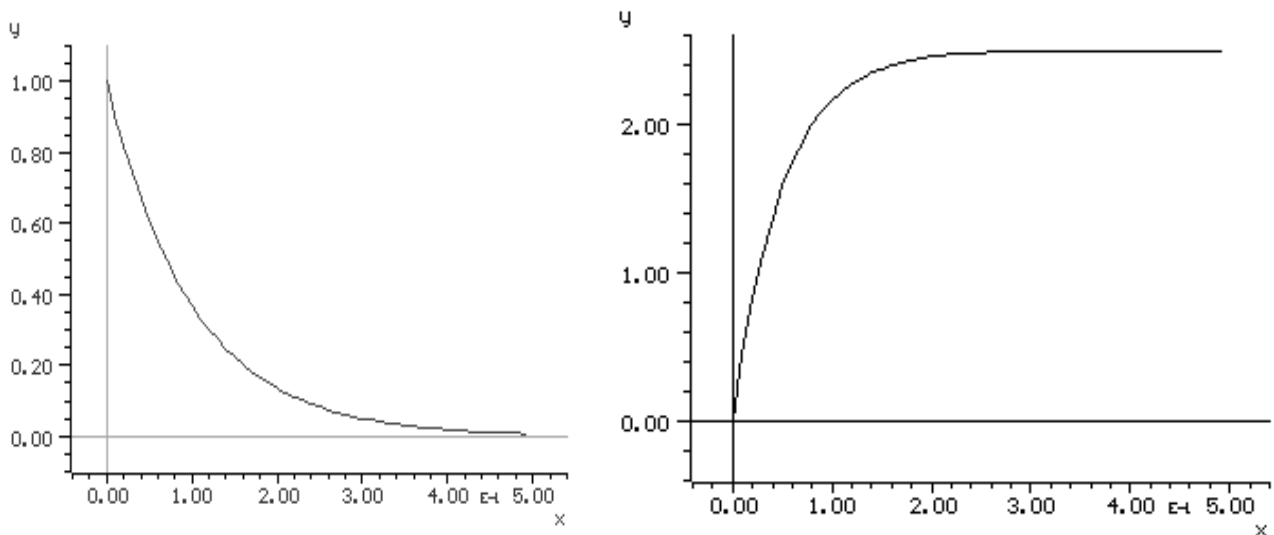
$$m_x(t) = x_0 e^{-\alpha t} + \alpha e^{-\alpha t} \int_0^t e^{\alpha\tau} m[n(\tau)] d\tau = x_0 e^{-\alpha t},$$

бо $m[n(t)] = 0$, а дисперсія:

$$\begin{aligned}
\sigma_x^2(t) &= m[(x(t) - m_x(t))^2] = m \left[\left(x_0 e^{-\alpha t} + \alpha e^{-\alpha t} \int_0^t e^{\alpha \tau} n(\tau) d\tau - x_0 e^{-\alpha t} \right) \times \right. \\
&\times \left. \left(x_0 e^{-\alpha t} + \alpha e^{-\alpha t} \int_0^t e^{\alpha s} n(s) ds - x_0 e^{-\alpha t} \right) \right] = \alpha^2 e^{-2\alpha t} \int_0^t \int_0^t e^{\alpha \tau} e^{\alpha s} m[n(\tau)n(s)] d\tau ds = \\
&= \alpha^2 e^{-2\alpha t} \frac{N_0}{2} \int_0^t \int_0^t e^{\alpha \tau} e^{\alpha s} \delta(\tau - s) d\tau ds = \alpha^2 e^{-2\alpha t} \frac{N_0}{2} \int_0^t e^{2\alpha \tau} d\tau = \frac{N_0 \alpha}{4} (1 - e^{-2\alpha t}).
\end{aligned}$$

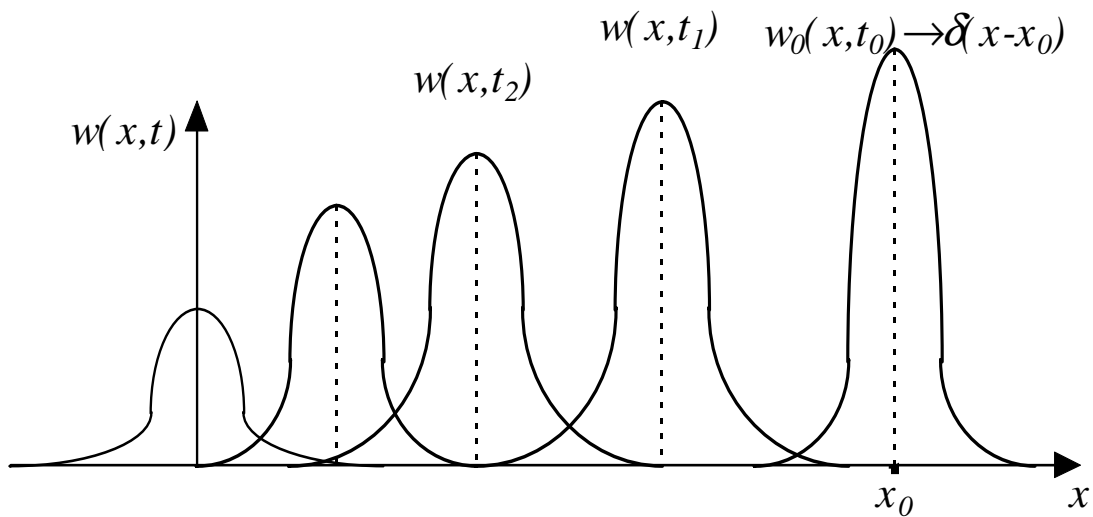
(використано “селективну” властивість δ -функції).

Графіки залежностей математичного сподівання і дисперсії від часу показано на мал. 4.3 (а, б) (при $R=1$ кОм, $C=0.1$ мкФ, $\alpha=10$, $N_0=1$).



Мал. 4.3.

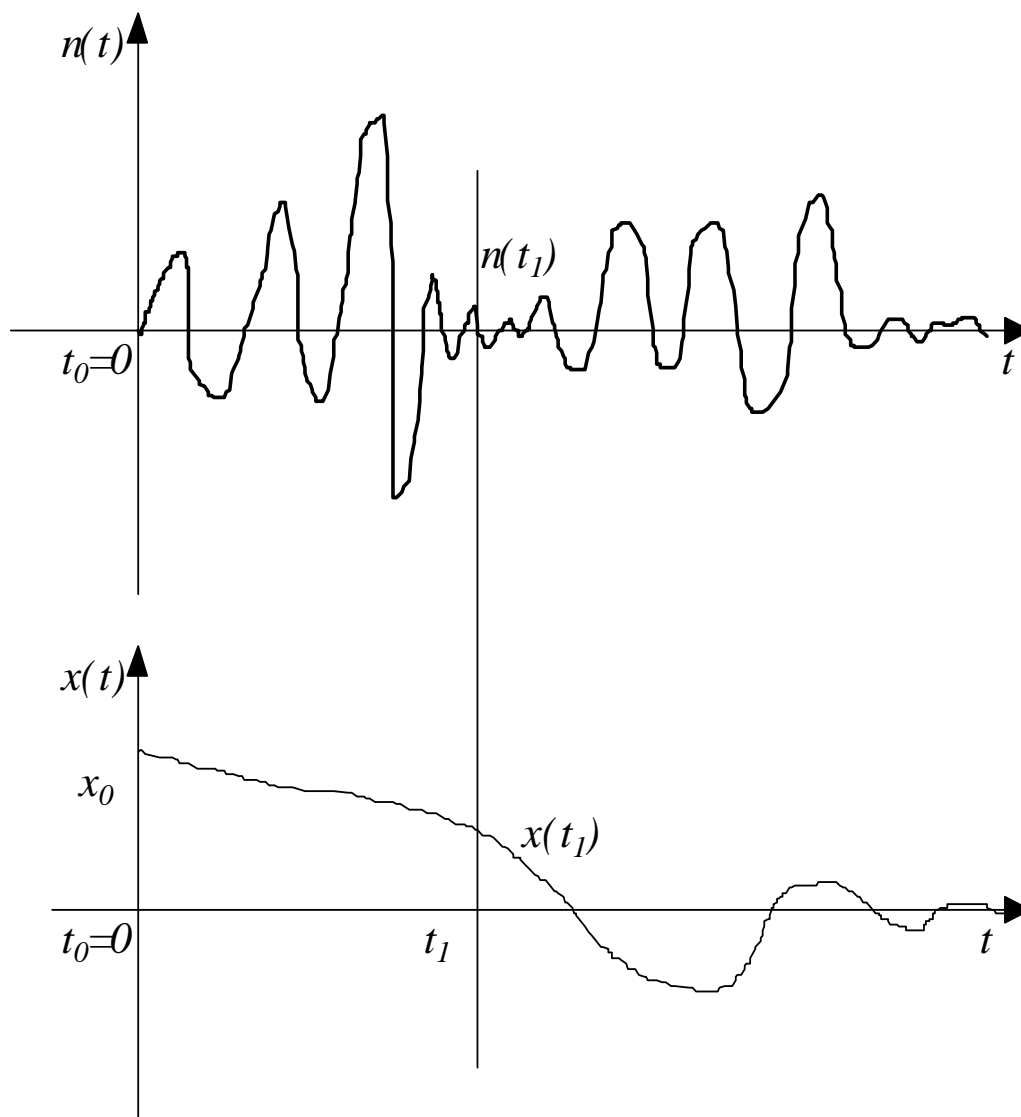
На мал. 4.4 показано зміну вигляду розподілу у часі.



Мал. 4.4. Зміни у часі вигляду закону розподілу імовірностей значень процесу $x(t)$ на виході RC -ланки

Коли $m_x(t) = 0$, а $\sigma^2 = \alpha \frac{N_0}{4}$, то процес стає стаціонарним, його густина розподілу від часу вже не залежить.

Вигляд процесів на вході і виході RC -ланки показано на мал. 4.5.



Мал. 4.5.

Виберемо момент часу $t = t_1 > t_0$ і зафіксуємо значення $x(t_1)$. Тоді процес $x(t)$ при $t > t_1$ буде визначатися виразом:

$$x(t) = e^{-\alpha(t-t_1)} \left[x(t_1) + \alpha \int_{t_1}^t e^{\alpha\tau} n(\tau) d\tau \right],$$

тобто, зовсім не залежить від значень $x(\theta)$, $\theta < t_1$. Але зауважимо, що це можливо лише для білого шуму $n(\tau)$, інакше, для процесу $\xi(t)$ на інтервалі $(t_1, t_1 + \tau_k)$ залежали б від значень $\xi(\theta)$, $t_1 - \tau_k < \theta < t_1$.

Так що процес $x(t)$ також марковий, але його прирости взяті на неперетинних інтервалах вже не є марковими (не є незалежними).

4.4.3. Визначення коефіцієнтів рівняння Фокера-Планка-Колмогорова (ФПК)

Неоднорідні диференціальні рівняння, які описують вплив лінійних кіл, систем на випадковий процес (білий шум, вінерівський процес), коли неоднорідність, збурення є випадковим процесом, називають стохастичними. Коли ці рівняння від ϑ , то можна визначити моменти функцій розподілу значень процесу на виході лінійного кола. Для марківських процесів з неперервним часом можна вивести диференціальне рівняння, невідомою функцією якого є функція розподілу імовірностей значень процесу. Це рівняння у частинних похідних, бо ця функція залежить від часу. Наприклад, рух броунівської частинки, процеси дифузії, теплопровідності описує рівняння в частинних похідних другого порядку параболічного вигляду. Для одновимірного випадку матимемо:

$$\frac{\partial \vartheta}{\partial t} = -K_1 \frac{\partial \vartheta}{\partial x} + \frac{K_2}{2} \frac{\partial^2 \vartheta}{\partial x^2},$$

де ϑ — густина імовірності переходу частинки від координати до координати, K_1 — коефіцієнт зносу, K_2 — коефіцієнт дифузії. Це рівняння називають ще дифузійним. Для ненулевих початкових умов x_0 і t_0 це рівняння має розв'язок:

$$\vartheta(x, t | x_0, t_0) = \frac{1}{\sqrt{2\pi K_2(t-t_0)}} \exp \left\{ -\frac{[x - x_0 - K_1(t-t_0)]^2}{2K_2(t-t_0)} \right\}.$$

Одновимірна густина імовірності знаходження частинки в координаті $x(t)$ відповідає такому ж рівнянню.

Коефіцієнти K_1 , K_2 шукаються за виразом, який отримується при виведенні рівняння:

$$K_n(x, t) = \lim_{\tau \rightarrow 0} \frac{m_n(x)}{\tau} = \lim_{\tau \rightarrow 0} \frac{m(x_\tau - x)^n}{\tau},$$

тобто,

$$K_1(x, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{\infty} (x_\tau - x) \vartheta(x_\tau, t + \tau | x, t) dx_\tau;$$

$$K_2(x, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{\infty} (x_\tau - x)^2 \vartheta(x_\tau, t + \tau | x, t) dx_\tau.$$

Тому, вибравши $t_0 = 0$, $x_0 = x$ для $t = \tau$ маємо:

$$x(\tau) = x_\tau = xe^{-\alpha\tau} + e^{-\alpha\tau} \alpha \int_0^\tau e^{\alpha s} n(s) ds,$$

а приріст —

$$x_\tau - x = x(e^{-\alpha\tau} - 1) + \alpha e^{-\alpha\tau} \int_0^\tau e^{\alpha s} n(s) ds.$$

Оскільки,

$$m[n(t)] = 0, \quad e^{-\alpha\tau} \approx 1 - \alpha\tau,$$

а

$$m[x_\tau - x] = x(e^{-\alpha\tau} - 1) = -\alpha x \tau,$$

(враховуючи, що для білого шуму перехідна густина імовірностей співпадає з одновимірною густиною), маємо і другий момент:

$$m[(x_\tau - x)^2] = x^2 \alpha^2 \tau^2 + \frac{N_0 \alpha^2 \tau}{2}.$$

У результаті отримаємо:

$$K_1(x, t) = K_1(x) = -\alpha x,$$

$$K_2(x, t) = K_2(x) = \alpha^2 \frac{N_0}{2}.$$

Так що для марківського процесу $x(t)$ на виході RC-ланки рівняння ФПК має вигляд:

$$\frac{\partial w(x,t)}{\partial t} = \alpha \frac{\partial}{\partial x} [xw(x,t)] + \frac{\alpha^2 N_0}{4} \frac{\partial^2 w(x,t)}{\partial x^2}.$$

Розв'язок цього рівняння є гаусова (нормальна) густина імовірності:

$$w(x,t) = \frac{1}{\sqrt{2\pi}\sigma_x(t)} \exp \left\{ -\frac{[x - m_x(t)]^2}{2\sigma_x^2(t)} \right\}.$$

Процес $x(t)$ з $K_1(x) = -\alpha x$, $K_2 = \text{const} = \alpha^2 \frac{N_0}{2}$ називають процесом Орнштейна-Уленбека.

Розглянувши розібраний тут випадок дії білого шуму на RC-ланку, легко дати фізичну інтерпретацію коефіцієнтів K_1 , K_2 .

Коефіцієнт K_1 виражає середню швидкість детермінованої зміни координати $x(t)$. А коефіцієнт K_2 виражає розкид цієї швидкості відносно її середнього значення. Таким чином, маємо можливість:

а) опису (моделювання) за допомогою стохастичного рівняння, коли збурення, неоднорідність є стохастичним процесом, а рівняння описує певний фізичний принцип, властивий змінам фізичної величини, який характеризує досліджуване явище (закони збереження, нерозривність і таке ін.);

б) за допомогою опису диференціального рівняння, у якому потрібною функцією є функція розподілу імовірностей значень цієї фізичної величини.

Інші уявлення про моделювання випадкових процесів у системах: лінійне зображення та спектральні (кореляційні).

Всі способи опису — за допомогою функцій розподілу, за допомогою моментів, кореляційна теорія дають свої вигоди при вирішенні проблем зв'язку, управління, контролю, ідентифікації та інше в різних галузях. Тут розглянуто "елемент" такого розв'язання, а саме — вплив ланки першого порядку на випадковий процес (який представлено білим шумом, вінерівським процесом, що мають властивості марковості). Можна зробити також висновок, що рівняння ФПК зв'язане з стохастичним рівнянням, тобто коефіцієнти рівняння ФПК можна шукати за виглядом стохастичного рівняння. Наприклад,

якщо $x'(t) = F(x, n(t))$, а $F(x, n(t)) = f(x, t) + g(t)n(t)$, то $K_1(x, t) = f(x, t)$, а $K_2(x, t) = \frac{N_0}{2} g^2(t)$.

Можна ставити обернену задачу — за коефіцієнтами дифузії та зносу шукати вигляд стохастичного рівняння. У загальному випадку цей перехід не однозначний. Але, наприклад, при вищезгаданому випадку $g(x) = \sqrt{K_2(x)}$, $f(x) = K_1(x) - \frac{1}{4} \frac{\partial K_2(x)}{\partial x}$, так що

$$x'(t) = K_1(x) - \frac{1}{4} \frac{\partial K_2(x)}{\partial x} + \sqrt{K_2(x)} n_0(t) \text{ (якщо } N_0 = 2 \text{)}.$$

Перехід від однієї моделі до іншої пов'язаний з необхідністю різноманітної фізичної інтерпретації. Так за допомогою рівняння ФПК описуємо процеси переносу речовин — конвекцією, дифузією (швидкістю конвекції і дифузії) як його механізмів. Звичайне диференціальне рівняння описує зміну значення величини в координатах і часі (цей же перенос, але з іншої точки зору — як інтегральну характеристику, без розгляду його механізму).

Підсумок

У загальному розумінні випадковий процес може бути заданий у термінах функцій розподілу або в рамках теорії 2-го порядку — спектрально-кореляційної. Тоді процес на виході лінійної ланки системи шукають у такому ж вигляді. У розділі розглянуто конкретні випадки, які мають практичну апробацію і застосування. Причому, лінійна ланка системи може задаватися у вигляді вхід-вихід (функцією передачі), диференціальним рівнянням, тощо. Задачу ж, яку вона розв’язує, трактують як згладжування, фільтрацію та ін. У наступних розділах розглядається задача обробки, коли її результатом є також процес, а не його характеристика — оптимальна фільтрація форми сигналів.

Методи “сильної” теорії (теорії вищого порядку, не гаусовий випадок, тощо) тут не розглядаються. Крім того в даному розділі розглянуто аналіз лінійної ланки, тобто, випадок, коли задано її структуру і параметри елементів (значення номіналів резисторів, конденсаторів, тощо, а також вигляд їх заєднання).

У наступних розділах розглянуто деякі випадки і методи синтезу, проте, реалізація (які елементи і їх заєднання) лінійної ланки системи не розглядається (за математичними операціями у виразах (формулах), якими подається розв’язок задачі синтезу, можна побудувати алгоритми чисельної реалізації, “схемна” ж реалізація не є такою видимою при розгляді формул).

Розділ 5

ОПТИМАЛЬНА ФІЛЬТРАЦІЯ

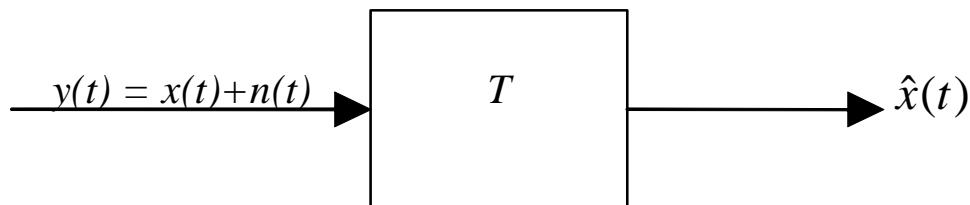
5.1. Критерій оптимальності

Досить часто адекватною до біосигналу є сума функцій $y(t)$ — детермінованої неперервної функції $x(t)$, інтегрованої „з квадратом” ($\int |x(t)|^2 dt < \infty$), та номального білого шуму $n(t)$:

$$y(t) = x(t) + n(t). \quad ()$$

Тоді обробка (опрацювання) біосигналу полягає у визначенні оцінки $\hat{x}(t)$ функції $x(t)$ за відомими шумом $n(t)$ та сумою $y(t)$. (В загальному випадку $x(t)$ і $n(t)$ є реалізаціями ергодичних випадкових функцій (випадкових процесів)). Ланка лінійної системи, яка виконує таку обробку, називається фільтром. Математичною моделлю такої ланки є лінійний оператор T (мал. 5.1):

$$\hat{x}(t) = T\{y(t)\}.$$



Мал. 5.1.

Фактично, корисний сигнал, що міститься у біосигналі, має вигляд якоїсь функції $x_d(t)$. Тому існує похибка моделювання $\xi_m(t) = x(t) - x_d(t)$, одна зі складових похибки фільтра. Іншою складовою є похибка фільтрації $\xi_f(t) = \hat{x}(t) - (x_d(t) + \xi_m(t))$, тобто, сумарна похибка фільтра

$$\xi(t) = \xi_m(t) + \xi_f(t) = \hat{x}(t) - x_d(t) \quad ()$$

Структура похибки адекватна до структури виразу (), якщо відповідні її складові є нормальним білим шумом. Очевидно, що ця похибка не може бути рівною нулеві. Але, серед усіх можливих характеристик фільтрів є така, для якої похибка буде мінімальною. Пошук такої характеристики називається оптимізацією фільтру. За критерій оптимальності обґрунтовується вибір середнього квадрату похибки фільтрації, оцінка сигналу є оптимальною, коли її значення мінімальне

$$I = M |\xi(t)|^2 \Big|_{min} = M |\hat{x}(t) - x(t)|^2 \Big|_{min}, \quad (5.1.1)$$

де M — символ математичного сподівання. Коли функції $x(t)$ і $n(t)$ ергодичні стаціонарні випадкові процеси з відомими спектральними густинами, то фільтр є лінійним, стаціонарним оператором з ядром $h(t)$

$$\hat{x}(t) = T\{y(t)\} = \int_{-\infty}^{\infty} h(\tau) y(t - \tau) d\tau$$

Цей оператор є фізично реалізованим лінійною стаціонарною ланкою з імпульсною ваговою функцією $h(t)$.

Формулювання задачі оптимізації оптимального фільтру набуває такого виразу

$$\arg \min_{h(\tau) \in C \subset L^2} M \left| \int_{-\infty}^{\infty} h(\tau) y(t - \tau) d\tau - x(t) \right|^2, \quad (5.1.3)$$

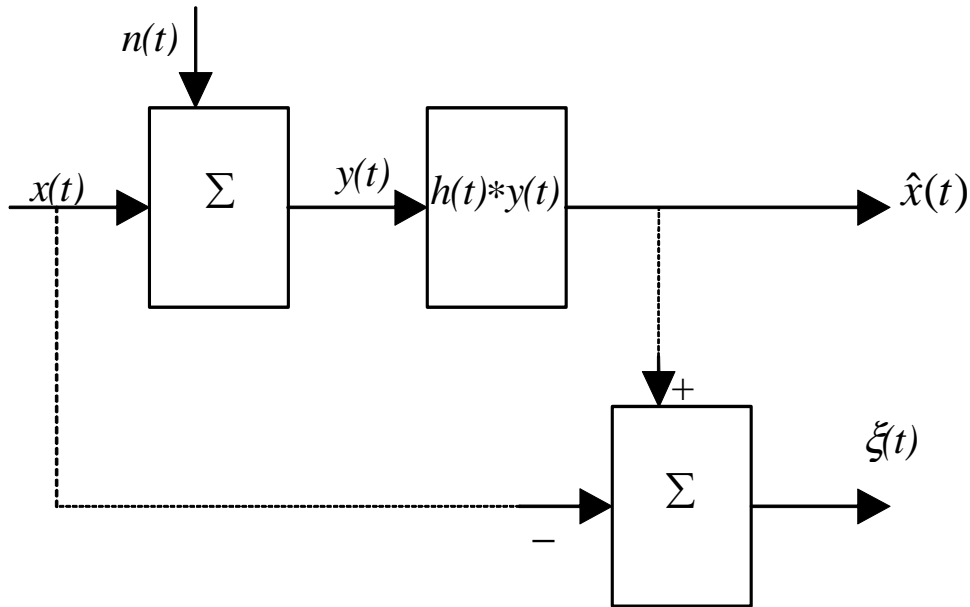
що означає: вибрати з множини C неперервних функцій, що належать функційному простору L^2 інтегрованих з квадратом, функцію h , яка забезпечує мінімальне значення математичного сподівання квадрату модуля різниці ().

Фільтр з імпульсною ваговою функцією h , визначеною за виразом (5.1.3), називається **фільтром Колмогорова-Вінера** або **оптимальним фільтром***.

* Про історію виникнення цієї назви див., наприклад, в книзі Н. Винер „Кибернетика”...

5.2. Схема завдання для побудови оптимального фільтру

З виразу (5.1.3) випливає схема (мал. 5.2), у якій наведено зв'язки поміж величинами, які зустрічаються у завданні оптимальної фільтрації.



Мал. 5.2.

На схемі мал. 5.2 наведено функції та оператори: $x(t)$ — корисний сигнал біооб'єкту; $n(t)$ — результуючий шум від джерел біологічного походження, відбору, зовнішніх чинників; адитивна суміш корисного сигналу і шуму, що поступає на вхід фільтру; $\hat{x}(t)$ — оцінка корисного сигналу (результат оптимальної фільтрації), $*$ — оператор згортки: $\int_{-\infty}^{\infty} h(\tau)y(t-\tau)d\tau$; $\xi(t) = \hat{x}(t) - x(t)$ — похибка фільтрації. Залишається невідомою $h(t)$ — імпульсна функція фільтру. Визначити функцію при заданих деяких умовах можна, наприклад, одним з методів варіаційного числення.

5.3. Визначення імпульсної функції оптимального фільтру

5.3.1. Частинна задача. Якщо для виразу (5.1.1)

$$I = M\xi^2(t) = M|\hat{x}(t) - x(t)|^2 \quad ()$$

математичні сподівання сигналу і шуму $m_x = m_n = 0$, то $m_y = 0$, $m_{\hat{x}} = 0$. Врахуємо ці умови, виконавши відповідні підстановки значень математичних сподівань в ():

$$\begin{aligned} I &= M[\hat{x}^2(t) - 2\hat{x}(t)x(t) + x^2(t)] = M\hat{x}^2(t) - 2M\hat{x}(t)x(t) + Mx^2(t) = \\ &= M\left[\int_{-\infty}^{\infty} h(\tau)y(t-\tau)d\tau \int_{-\infty}^{\infty} h(t-\theta)y(\theta)d\theta\right] - 2M\left[x(t) \int_{-\infty}^{\infty} h(t-\theta)y(\theta)d\theta\right] + Mx(t)x(t) = \\ &= M\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y(\tau)y(\theta)h(t-\tau)h(t-\theta)d\tau d\theta\right] - 2M\int_{-\infty}^{\infty} x(t)h(t-\theta)y(\theta)d\theta + Mx(t)x(t) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} My(\tau)y(\theta)h(t-\tau)h(t-\theta)d\tau d\theta - 2\int_{-\infty}^{\infty} Mx(t)h(t-\theta)y(\theta)d\theta + Mx(t)x(t) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta)h(t-\tau)h(t-\theta)d\tau d\theta - 2\int_{-\infty}^{\infty} R_{xy}(t, \theta)h(t-\theta)d\theta + R_{xx}(t, t) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta)h(t, \tau)h(t, \theta)d\tau d\theta - 2\int_{-\infty}^{\infty} R_{xy}(t, \theta)h(t, \theta)d\theta + R_{xx}(t, t), \end{aligned}$$

де $R_{yy}(\cdot)$, $R_{xy}(\cdot)$ — автокореляційна та взаємнокореляційна функції. Отже,

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta)h(t, \tau)h(t, \theta)d\tau d\theta - 2\int_{-\infty}^{\infty} R_{xy}(t, \theta)h(t, \theta)d\theta + R_{xx}(t, t), \quad (5.3.1)$$

тобто, I є сумою квадратичного та лінійного функціоналів від $h(t, \theta)$ і „постійної” функції, причому $h(t, \theta)$ є функцією від θ при фіксованому (параметрі) t . За означенням, $h(\cdot, \cdot)$ буде оптимальною, коли I набире мінімального значення. Умовою мінімуму функціоналу є вираз $grad I = \partial I / \partial \theta = 0$, звідки отримаємо рівняння:

$$\int_{-\infty}^{\infty} R_{yy}(\tau, \theta)h(t, \tau)d\tau = R_{xy}(t, \theta). \quad (5.3.2)$$

Оскільки для стаціонарних процесів

$$R_{yy}(\tau, \theta) = R_{yy}(\tau - \theta), h(t, \tau) = h(t - \tau), R_{xy}(t, \theta) = R_{xy}(t - \theta),$$

то (5.3.2) набирає вигляду

$$\int_{-\infty}^{\infty} R_{yy}(\tau - \theta) h(t - \tau) d\tau = R_{xy}(t - \theta),$$

або, пере позначивши аргументи $t - \theta = s$, $t - \tau = \sigma$, $\tau - \theta = s - \sigma$

$$\int_{-\infty}^{\infty} R_{yy}(s - \sigma) h(\sigma) d\sigma = R_{xy}(s). \quad (5.3.3)$$

Вираз (5.3.3) називається інтегральним *рівнянням Вінера-Хопфа*^{*}. Постає задача: за відомими авто кореляційною R_{yy} та взаємно кореляційною R_{xy} функціями, відповідно відібраного сигналу y та відібраного сигналу y і сигналу x , знайти з рівняння (5.3.3) імпульсну функцію h таку, при якій буде справедливою умова (5.1.3).

З врахуванням умови фізичної реалізованості $h(t) = 0$, при $t < 0$ рівняння Вінера-Хопфа набире вигляду:

$$\int_0^{\infty} R_{yy}(s - \sigma) h(\sigma) d\sigma = R_{xy}(s) \quad (5.3.4)$$

5.3.2. Задача варіаційного числення^{}.** Інший спосіб одержання рівняння для визначення імпульсної функції оптимального фільтру (базований на застосуванні поняття варіації функції) не потребує накладання умов на математичні сподівання. Нехай $h_0(t)$ — функція, для якої () набирає мінімального значення. Постає запитання: чи існує ще одна така функція $h(t)$? Для відповіді на це запитання виразимо цю нову функцію шляхом надання значенням

* Про походження цієї назви див., наприклад, в книзі.....

**

функції $h_0(t)$ деяких приростів (тобто, надамо значенням функції змін, варіації)

$$h(t)=h_0(t)+ag(t), \quad ()$$

де, a — не залежить від t , а $g(t)$ — деяка функція, $g(t) \neq 0$. За означенням, при $a \neq 0$ величина I більша, ніж при $a = 0$, тому:

$$\left. \frac{dI}{da} \right|_{a=0} = 0 \quad (5.3.5)$$

Підставимо $h(t)=h_0(t)+ag(t)$, у вираз для I :

$$\begin{aligned} I = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta) [h_0(t, \tau) + ag(t, \tau)] [h_0(t, \theta) + ag(t, \theta)] d\tau d\theta - \\ & - 2 \int_{-\infty}^{\infty} R_{xy}(t, \theta) [h_0(t, \theta) + ag(t, \theta)] d\theta + R_{xx}(t, t). \end{aligned}$$

Тоді

$$\begin{aligned} \frac{dI}{da} = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta) h_0(t, \tau) g(t, \theta) d\tau d\theta + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta) g(t, \tau) h_0(t, \theta) d\tau d\theta + \\ & + 2a \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta) g(t, \tau) g(t, \theta) d\tau d\theta - 2 \int_{-\infty}^{\infty} R_{xy}(t, \theta) g(t, \theta) d\theta = \\ = & 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta) h_0(t, \tau) g(t, \theta) d\tau d\theta + 2a \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(\tau, \theta) g(t, \tau) g(t, \theta) d\tau d\theta - \\ & - 2 \int_{-\infty}^{\infty} R_{xy}(t, \theta) g(t, \theta) d\theta. \end{aligned}$$

Поклавши $a = 0$, отримаємо вираз:

$$\int_{-\infty}^{\infty} g(t, \theta) \left[\int_{-\infty}^{\infty} R_{yy}(\tau, \theta) h_0(t, \tau) d\tau - R_{xy}(t, \theta) \right] d\theta = 0. \quad ()$$

Оскільки $g(t) \neq 0$, то

$$\int_{-\infty}^{\infty} R_{yy}(\tau, \theta) h_0(t, \tau) d\tau - R_{xy}(t, \theta) = 0,$$

тобто отримано потрібне рівняння. Для стаціонарних сигналів:

$$\int_{-\infty}^{\infty} R_{yy}(\tau - \theta) h_0(t - \tau) d\tau = R_{xy}(t - \theta).$$

Замінивши змінні $t - \theta = s$, $t - \tau = \sigma$, $\tau - \theta = s - \sigma$, отримаємо рівняння

Вінера-Хопфа $\int_{-\infty}^{\infty} R_{yy}(s - \sigma) h_0(\sigma) d\sigma = R_{xy}(s)$, яке, при врахуванні умови

фізичної реалізованості, набуває вигляду: $\int_0^{\infty} R_{yy}(s - \sigma) h_0(\sigma) d\sigma = R_{xy}(s)$.

Отже, для лінійної, стаціонарної, фізично реалізованої ланки (системи) існує єдина імпульсна функція, яка забезпечує виконання оптимальної за критерієм () фільтрації (зворотнє твердження несправедливе — ланок, систем з такою, оптимальною імпульсною функцією є безліч; для побудови, синтезу ланки знання її імпульсної функції недостатньо).

Отримання виразу подібного до (), з добутком функцій під інтегралом, одна з яких завідомо рівна нулеві, є типовим для розв'язування задач методом варіаційного числення (як і обґрунтований на евристичних засадах вибір виразу вигляду ()).

5.4. Отримання рівняння Вінера-Хопфа методом ортогоналізації

Якщо T — лінійний оператор, адекватний як математична модель оптимальному фільтру Колмогорова-Вінера, тобто такий, що I набуває мінімального значення, то різниця $T\{y(t)\} - x(t)$ ортогональна до $y(t)$ при кожному значенні t . Тоді

$$M[(T\{y(t)\} - x(t))y(t)] = 0, t \in (-\infty, \infty). \quad (5.4.1)$$

Нехай існує T_0 — деякий інший такий лінійний оператор. Тоді (для зручності спрощено позначення):

$$\begin{aligned}
I &= M[T_0 y - x]^2 = M[(T_0 y - Ty) + (Ty - x)]^2 = \\
&= M[(T_0 - T)y]^2 + 2(T_0 - T)y(Ty - x) + (Ty - x)^2 = \\
&= M[(T_0 - T)y]^2 + 2M(T_0 - T)y(Ty - x) + M(Ty - x)^2.
\end{aligned}$$

Якщо (5.4.1) справедливо, то другий доданок рівний нулю. Перший — не від'ємний. Таким чином, оператор T_0 не забезпечить меншої середньоквадратичної похибки, ніж оператор T , тому $T_0 = T$.

Вираз (5.4.1) є представленням так званого **принципу ортогональності**. З врахуванням того, що

$$Ty(t) = \hat{x}(t) = \int_{-\infty}^{\infty} h(t - \tau)y(\tau)d\tau,$$

принцип ортогональності можна записати так:

$$M\left[\left[\int_{-\infty}^{\infty} h(t - \tau)y(\tau)d\tau - x(t)\right]y(t)\right] = 0,$$

звідки:

$$\begin{aligned}
\int_{-\infty}^{\infty} My(\tau)y(s)h(t - \tau)d\tau &= Mx(t)y(s), \\
\int_{-\infty}^{\infty} R_{yy}(\tau, s)h(t - \tau)d\tau &= R_{xy}(t, s).
\end{aligned}$$

За умови стаціонарності

$$\int_0^{\infty} R_{yy}(\tau - s)h(t - \tau)d\tau = R_{xy}(t - s).$$

Заміною змінних $t - \theta = s$, $t - \tau = \sigma$, $\tau - \theta = s - \sigma$ отримаємо:

$$\int_{-\infty}^{\infty} R_{yy}(\theta - \sigma)h(\sigma)d\sigma = R_{xy}(\theta),$$

рівняння Вінера-Хопфа. Якщо врахувати умову фізичної реалізованості (тобто $h(t) = 0$ при $t < 0$), це рівняння набирає такого вигляду:

$$\int_0^{\infty} R_{yy}(\theta - \sigma)h(\sigma)d\sigma = R_{xy}(\theta).$$

5.5. Розв'язування рівняння Вінера-Хопфа

Оптимальна фільтрація полягає у лінійній обробці сигналу імпульсною функцією фільтру, такою, що середній квадрат похибки на виході фільтру мінімальний. Пошук такої імпульсної функції зводиться до розв'язування рівняння Вінера-Хопфа —

$$\int_{-\infty}^{\infty} R_{yy}(s - \sigma)h(\sigma)d\sigma = R_{xy}(s). \quad (5.5.1)$$

Розв'язок даного рівняння легко знайти, коли виконати перетворення Лапласа від лівої і правої частини виразу (5.5.1):

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(s - \sigma)h(\sigma)e^{-ius}d\sigma ds = \int_{-\infty}^{\infty} R_{xy}(s)e^{-ius}ds.$$

Зробимо деякі зміни в лівій частині:

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{yy}(s - \sigma)h(\sigma)e^{-iu(s-\sigma)}e^{-iu\sigma}d\sigma ds &= \int_{-\infty}^{\infty} R_{xy}(s)e^{-ius}ds, \\ \int_{-\infty}^{\infty} R_{yy}(s - \sigma)e^{-iu(s-\sigma)}ds \int_{-\infty}^{\infty} h(\sigma)e^{-iu\sigma}d\sigma &= \int_{-\infty}^{\infty} R_{xy}(s)e^{-ius}ds, \end{aligned}$$

або:

$$F_{yy}(iu)H(iu) = F_{xy}(iu),$$

де $F_{yy}(iu)$ — спектральна густина потужності $y(t)$, $F_{xy}(iu)$ — взаємна спектральна густина потужності $x(t)$ і $y(t)$, $H(iu)$ — функція передачі фільтру. Отже, функція передачі оптимального фільтру:

$$H(iu) = \frac{F_{xy}(iu)}{F_{yy}(iu)}.$$

Використовуючи зворотнє перетворення Лапласа від $H(iu)$, можна знайти оптимальну імпульсну функцію $h(t)$.

Розглянемо випадок, коли корисний сигнал $x(t)$ і шум $n(t)$ некорельовані, тобто $F_{xn}(iu) = 0$. Тоді $F_{yy}(iu) = F_{xx}(iu) + F_{nn}(iu) = F_x(iu) + F_n(iu)$, $F_{xy}(iu) = F_{xx}(iu) = F_x(iu)$, тоді оптимальна функція передачі:

$$H(iu) = \frac{F_x(iu)}{F_x(iu) + F_n(iu)}.$$

5.6. Похибка оптимальної фільтрації

Знайдемо складову ξ_f похибки оптимального фільтра, яка визначається значенням мінімуму середнього квадрату відхилення

$$I_{min} = M\xi_f^2 = M|\hat{x} - x|^2$$

оцінки \hat{x} сигналу від сигналу x . Якщо T — лінійний оператор (математична модель оптимального фільтру), то $\hat{x} = Ty$, тоді

$$I_{min} = M|Ty - x|^2 = M[(Ty - x)(Ty - x)] = M[(Ty - x)Ty - (Ty - x)x].$$

Доданок $(Ty - x)Ty = \xi\hat{x} = 0$, оскільки $y = \hat{x} + \xi$ (принцип ортогональності похибки до оцінки). Отже,

$$I_{min} = M[(x - Ty)x]. \quad (5.6.1)$$

Оскільки $T\{y(t)\} = \int_{-\infty}^{\infty} y(t)h(t - \tau)d\tau$, і якщо $m_x = m_y = 0$, то:

$$I_{\min} = M[(x(t) - \int_{-\infty}^{\infty} y(\tau)h(t-\tau)d\tau)x(\theta)] = R_{xx}(t, \theta) - \int_{-\infty}^{\infty} R_{xy}(\tau, \theta)h(t-\tau)d\tau$$

і

При умові стаціонарності

$$I_{\min} = R_{xx}(t - \theta) - \int_{-\infty}^{\infty} R_{xy}(\tau - \theta)h(t - \tau)d\tau.$$

Замінивши змінні $t - \theta = s$, $t - \tau = \sigma$, $\tau - \theta = s - \sigma$, отримаємо:

$$I_{\min} = R_{xx}(s) - \int_{-\infty}^{\infty} R_{xy}(s - \sigma)h(\sigma)d\sigma.$$

Перетворенням Фур'є від лівої і правої частин останнього виразу

$$\begin{aligned} F\{I_{\min}\} &= \int_{-\infty}^{\infty} R_{xx}(s)e^{-ius}ds - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{xy}(s - \sigma)h(\sigma)e^{-ius}dsd\sigma = \\ &= \int_{-\infty}^{\infty} R_{xx}(s)e^{-ius}ds - \int_{-\infty}^{\infty} R_{xy}(s - \sigma)e^{-iu(s-\sigma)}ds \int_{-\infty}^{\infty} h(\sigma)e^{-iu\sigma}d\sigma = \\ &= F_{xx}(iu) - F_{xy}(iu)H(iu), \end{aligned}$$

де $F_{xx}(iu)$, $F_{xy}(iu)$ — спектральні густини потужності, $H(iu)$ — функція передачі фільтру, отримаємо спектральні представлення похибки.

При зворотньому перетворенні Фур'є —

$$I_{\min} = \int_{-\infty}^{\infty} [F_{xx}(iu) - F_{xy}(iu)H(iu)]e^{iut}du,$$

або, з врахуванням того, що $H(iu) = \frac{F_{xy}(iu)}{F_{yy}(iu)}$,

$$I_{\min} = \int_{-\infty}^{\infty} \frac{F_{xx}(iu)F_{yy}(iu) - |F_{xy}(iu)|^2}{F_{yy}(iu)}e^{iut}du. \quad (5.6.2)$$

Формула (5.6.1) пояснює зміст середнього квадрату похибки в просторі випадкових величин, а формула (5.6.2) дає змогу знайти його значення.

Підсумок

Очевидно, що ідея екстремуму використовується і для синтезу характеристик інтерполяторів, прогнозаторів, згладжувачів і т. ін. Зовнішній вигляд реалізації може бути означений критерієм.

У цьому розділі показано, як з мінімальною середньоквадратичною похибкою можна виділити сигнал з суміші його з шумом, вказано умови, за яких така модель є адекватною. Отже, крім моделювання зображення (представлення) сигналу у вигляді спектрів (розподілу потужності по частотах) та аналізу впливу на нього лінійних перетворень — знаходження спектрів на виході лінійної ланки (у випадку стаціонарних сигналів) чи знаходженні функцій розподілу на виході лінійної ланки при заданій вхідній функції розподілу імовірностей значень процесу можна ставити і розв'язувати задачі синтезу (пошуку) характеристик самої ланки при певних вимогах до характеристик синтезу на її виході при заданому вхідному сигналові. За контекстом розділу зрозуміло, що означають слова “заданий”, “певні вимоги” і т. п. Методи реалізації отриманої характеристики лінійною ланкою тут не розглянуто. Проте, зауважимо, що є труднощі (проблеми) реалізації. Вони лежать в області засобів і методів реалізації і виглядають як інструментальні похибки - основні, методичні, динамічні, впливу і причиною мають певну фізичну природу засобів. Буває так, що математично коректну характеристику фізично реалізувати не вдається із-за властивостей (можливостей) засобів її реалізації.

Основною ідеєю, на якій базується синтез, є пошук такої характеристики (функції), при якій досягається екстремум критерію (функціоналу), побудованого, вибраного як математичний об'єкт з фізичних уявлень про якість функціонування ланки.

Такі ланки (системи) називають екстремальними (деколи додають ще назву характеристики — екстремальні кореляційні, спектральні, тощо). Тепер більш поширеною назвою є назва оптимальні (фільтри, прогнозатори, тощо).

Подібну задачу (історично раніше за Вінера, що стало відомо пізніше через умови закритості робіт) розв'язав Колмогоров (не в рамках кореляційної теорії, а функцій розподілу).

Розділ 6

ФІЛЬТР КАЛМАНА-Б'ЮСІ

6.1. Задача Калманівської фільтрації

У попередньому розділі був розглянутий фільтр Вінера. Він являє собою лінійну ланку, на вхід якої надходить суміш корисного сигналу і шуму, а з виходу відбирається сигнал, що максимально наближається до корисного. Ми шукали імпульсну вагову функцію такого фільтру.

Проте, в багатьох випадках система може мати не один, а декілька входів, виходів (спостережних станів), а також неспостережні стани. Такі системи зручно розглядати у просторі станів.

Розглянемо таку лінійну систему. Через $x(t)$ позначимо її вектор станів. Припустимо, що модель зміни вектору станів задана системою лінійних диференціальних рівнянь —

$$\dot{x}(t) = F(t)x(t) + G(t)w(t), \quad (6.1.1)$$

у якому вектор $w(t)$ є білим шумом з нульовим середнім значенням, а елементи матриць $F(t)$ і $G(t)$ — відомі, у загальному випадку неперервні функції часу. Вектори $x(t)$ і $\dot{x}(t)$ мають розмір $n \times 1$, вектор $w(t)$ — $r \times 1$, матриця $F(t)$ — $n \times n$ і матриця $G(t)$ — $n \times r$ розміру.

Досить часто вектор станів $x(t)$ ми безпосередньо спостерігати не можемо, а замість нього спостерігаємо вектор $y(t)$, який одержуємо на виході системи —

$$y(t) = H(t)x(t) + v(t), \quad (6.1.2)$$

де $v(t)$ також є білим шумом з нульовим математичним сподіванням і матриця $H(t)$ — відома. Розміри величин, які входять в рівняння (6.1.2), наступні: вектори $y(t)$ і $v(t)$ — $m \times 1$, матриця $H(t)$ — $m \times n$.

Вважаємо, що початкове значення $x(t_0)$ вектора станів являє собою випадковий вектор з математичним сподіванням:

$$M\{x(t_0)\} = \overline{x_0}$$

і коваріаційною матрицею:

$$M\{x(t_0)x^T(t_0)\} = P_0.$$

Тут $x^T(t_0)$ означає транспонований вектор.

Величини $w(t)$ і $v(t)$ називаються, відповідно, вхідним шумом і шумом спостереження. Ми не робитимемо жодних спеціальних зауважень щодо функції розподілу вхідного шуму та шуму спостереження, ані щодо вектора початкового стану системи. Відзначимо лише, що для векторів $w(t)$ і $v(t)$, крім нульового математичного сподівання, задані коваріаційні матриці:

$$M\{w(t)\} = 0, M\{w(t)w^T(\tau)\} = Q(t)\delta(t - \tau);$$

$$M\{v(t)\} = 0, M\{v(t)v^T(\tau)\} = R(t)\delta(t - \tau),$$

де $\delta(t - \tau)$ — дельта-функція, а $Q(t)$, $R(t)$ — відомі квадратні матриці.

Нехай, окрім того, вектор станів, вхідний шум і шум спостереження взаємно не корельовані —

$$M\{x(t)v^T(\tau)\} = 0;$$

$$M\{x(t)w^T(\tau)\} = 0;$$

$$M\{v(t)w^T(\tau)\} = 0.$$

Оскільки ми не можемо безпосередньо спостерігати вектор станів, задача полягає в знаходженні лінійної незміщеної оцінки вектора $x(t)$ на основі результатів вимірів $y(t)$. Ця оцінка позначається через $\hat{x}(t)$. Похибка оцінювання:

$$\tilde{x}(t) = x(t) - \hat{x}(t).$$

Матриця коваріації похибки у випадку незміщеної оцінки рівна:

$$\tilde{P}(t) = M\{\tilde{x}(t)\tilde{x}^T(t)\}.$$

Оцінка повинна бути оптимальною в тому розумінні, що компоненти похибки оцінювання повинні мати мінімальну дисперсію. Ця умова еквівалентна вимозі

$$\text{tr}[\tilde{P}(t)] = \min,$$

де $\text{tr}[\tilde{P}(t)]$ — слід матриці $\tilde{P}(t)$.

Поставлена вище задача називається задачею калманівської фільтрації.

6.2. Критерій оптимальності фільтру.

Перед нами стоїть задача: для системи, заданої матричними рівняннями:

$$\dot{x}(t) = F(t)x(t) + G(t)w(t),$$

$$y(t) = H(t)x(t) + v(t),$$

знайти лінійну незміщену оцінку $\hat{x}(t)$ вектора станів $x(t)$ по результатах спостережень $y(t)$.

Очевидно, для того, щоб фільтр був оптимальний, похибка оцінювання

$$\tilde{x}(t) = x(t) - \hat{x}(t)$$

повинна бути мінімальною. Оскільки $x(t)$ являє собою випадковий вектор, розглядають евклідову норму $|\tilde{x}|$ цього вектора. Позначимо через I математичне сподівання від $|\tilde{x}|^2$. У випадку n -вимірного вектора похибок можна записати:

$$I = M|\tilde{x}|^2 = M\{\tilde{x}_1^2 + \tilde{x}_2^2 + \dots + \tilde{x}_n^2\}$$

В якості критерію оптимальності фільтру Калмана-Б'юсі приймають умову

$$I = \min, \quad (6.2.1)$$

тобто середній квадрат похибки оцінювання повинен бути мінімальним. Часто говорять також про мінімальне значення дисперсії похибки.

Умову оптимальності фільтру можна записати по-іншому, використавши матричне числення. Коваріаційна матриця похибки

$$\tilde{P}(t) = M \{ \tilde{x}(t) \tilde{x}^T(t) \}$$

являє собою квадратну матрицю розміру $n \times n$, по головній діагоналі якої розміщені елементи $M\tilde{x}_1^2, M\tilde{x}_2^2, \dots, M\tilde{x}_n^2, \dots$. Як відомо, сума діагональних елементів будь-якої квадратної матриці A називається **слідом** цієї матриці і позначається $tr[A]$. З урахуванням цього критерій оптимальності (6.2.1) набуде вигляду:

$$tr[\tilde{P}(t)] = M \{ \tilde{x}_1^2 + \tilde{x}_2^2 + \dots + \tilde{x}_n^2 \} = \min. \quad (6.2.2)$$

Критерій (6.2.2) і використовується для розв'язку задачі Калманівської фільтрації.

6.3. Поновлюючий процес у фільтрі Калмана-Б'юсі

Будемо шукати лінійну незміщену оцінку вектора станів $x(t)$ системи, якщо система задана матричними рівняннями

$$\dot{x}(t) = F(t)x(t) + G(t)w(t), \quad (6.3.1)$$

$$y(t) = H(t)x(t) + v(t), \quad (6.3.2)$$

і вектор спостережень $y(t)$ відомий. Умови, які накладаються на величини, що входять в рівняння (6.3.1) і (6.3.2), описані в параграфі 6.1.

Припустимо, що оцінка $\hat{x}(t)$ вектора станів відповідає диференціальному рівнянню:

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + K(t)y(t), \quad (6.3.3)$$

де $y(t)$ — відомий вектор спостережень, $A(t)$ і $K(t)$ — невідомі матриці. Для того, щоб оцінка була незміщеною, повинна виконуватись рівність:

$$M\{\hat{x}(t)\} = M\{x(t)\} = \bar{x}(t). \quad (6.3.4)$$

Обчислимо математичне сподівання від обох частин рівняння (6.3.3)

$$M\{\dot{\hat{x}}(t)\} = A(t)M\{\hat{x}(t)\} + K(t)M\{y(t)\}. \quad (6.3.5)$$

Але з (6.3.2) випливає, що:

$$M\{y(t)\} = M\{H(t)x(t) + v(t)\} = H(t)M\{x(t)\}, \quad (6.3.6)$$

так як $M\{v(t)\} = 0$.

Як наслідок, на основі (6.3.4) — (6.3.6) одержуємо диференціальне рівняння для середнього значення вектору станів системи:

$$\dot{\bar{x}}(t) = [A(t) + K(t)H(t)]\bar{x}(t). \quad (6.3.7)$$

Обчислимо математичне сподівання від обох частин рівняння (6.3.1), одержимо ще одне рівняння для $\bar{x}(t)$ (враховуємо, що $M\{w(t)\} = 0$)

$$\dot{\bar{x}} = F(t)\bar{x}(t). \quad (6.3.8)$$

Порівнюючи два останні рівняння, можна виписати першу умову незміщеності оцінки вектора станів —

$$A(t) = F(t) - K(t)H(t). \quad (6.3.9)$$

Друга умова полягає в тому, щоб розв'язки рівнянь (6.3.7) і (6.3.8) відшукувались при одному і тому ж початковому значенні

$$\bar{x}(t_0) = M\{\hat{x}(t_0)\} = M\{x(t_0)\} = \bar{x}_0. \quad (6.3.10)$$

Якщо виконати умови незміщеності, користуючись рівностями (6.3.9) і (6.3.10), то алгоритм фільтру (6.3.3) набуде вигляду:

$$\begin{aligned} \dot{\hat{x}}(t) &= F(t)\hat{x}(t) + K(t)[y(t) - H(t)\hat{x}(t)], \\ \hat{x}(t_0) &= \bar{x}_0. \end{aligned} \quad (6.3.11)$$

Такий запис рівняння дозволяє інтерпретувати його наступним чином. Якщо матриця $K(t)$ є оптимальною, наприклад вибрана із умови забезпечення мінімального значення дисперсії похибки, то різниця $y(t) - H(t)\hat{x}(t)$ виявляється шумом, який називається **поновлюючим процесом**, оскільки саме ця різниця містить в собі всю нову інформацію, яка одержується після спостереження значень $y(t)$. Матрицю $K(t)$ називають матрицею коефіцієнтів фільтру.

6.4. Коефіцієнти фільтру Калмана-Б'юсі.

Знайдемо тепер матрицю коефіцієнтів $K(t)$, таку, щоб фільтр, який описується алгоритмом:

$$\hat{x}(t) = F(t)\hat{x}(t) + K(t)[y(t) - H(t)\hat{x}(t)], \quad (6.4.1)$$

забезпечував мінімальну дисперсію похибки оцінювання вектора станів системи, заданої матричними рівняннями:

$$\dot{x}(t) = F(t)x(t) + G(t)w(t), \quad (6.4.2)$$

$$y(t) = H(t)x(t) + v(t). \quad (6.4.3)$$

Тобто від $K(t)$ залежить мінімальне значення величини

$$I = \text{tr}[\tilde{P}(t)], \quad (6.4.4)$$

де $\tilde{P}(t)$ являє собою коваріаційну матрицю похибки, $\tilde{P}(t) = M\{\tilde{x}(t)\tilde{x}^T(t)\}$. Спочатку виведемо вираз для цієї матриці. Так як похибка оцінювання —

$$\tilde{x}(t) = x(t) - \hat{x}(t), \quad (6.4.5)$$

то, очевидно,

$$\dot{\tilde{x}}(t) = \dot{x}(t) - \dot{\hat{x}}(t).$$

Із (6.4.1) і (6.4.2) одержуємо:

$$\begin{aligned}\dot{\tilde{x}}(t) = \dot{x}(t) - \dot{\hat{x}}(t) = F(t)x(t) + G(t)w(t) - F(t)\hat{x}(t) - \\ - K(t)[y(t) - H(t)\hat{x}(t)].\end{aligned}$$

Якщо тепер скористатися представленням (6.4.3) і позначенням (6.4.5), то можна записати:

$$\dot{\tilde{x}}(t) = [F(t) - K(t)H(t)]\tilde{x}(t) + u(t), \quad (6.4.6)$$

де

$$u(t) = G(t)w(t) - K(t)v(t) \text{ —}$$

білий шум з нульовим математичним сподіванням і моментами

$$M\{u(t)u^T(\tau)\} = [G(t)Q(t)G^T(t) + K(t)R(t)K^T(t)]\delta(t - \tau),$$

$$M\{\tilde{x}(t)u^T(\tau)\} = 0.$$

Задачу зведено до дослідження матричного диференціального рівняння для коваріаційної матриці похибки. Знайдемо його. Виходячи з того, що $\tilde{P}(t) = M\{\tilde{x}(t)\tilde{x}^T(t)\}$, запишемо:

$$\dot{\tilde{P}}(t) = M\{\dot{\tilde{x}}(t)\tilde{x}^T(t) + \tilde{x}(t)\dot{\tilde{x}}^T(t)\} \quad (6.4.7)$$

Підставимо в рівняння (6.4.7) вираз (6.4.6), одержимо:

$$\begin{aligned}\dot{\tilde{P}}(t) = M\{[F(t) - K(t)H(t)]\tilde{x}(t)\tilde{x}^T(t) + u(t)\tilde{x}^T(t) + \\ + \tilde{x}(t)u^T(t)[F(t) - K(t)H(t)]^T + \tilde{x}(t)u^T(t)\} = \\ = [F(t) - K(t)H(t)]M\{\tilde{x}(t)\tilde{x}^T(t)\} + M\{\tilde{x}(t)\tilde{x}^T(t)\}[F(t) - K(t)H(t)]^T + \\ + M\{u(t)\tilde{x}^T(t) + \tilde{x}(t)u^T(t)\} = [F(t) - K(t)H(t)]\tilde{P}(t) + \\ + \tilde{P}(t)[F(t) - K(t)H(t)]^T + M\{u(t)\tilde{x}^T(t) + \tilde{x}(t)u^T(t)\}.\end{aligned}$$

Можна показати, що:

$$M\{u(t)\tilde{x}^T(t) + \tilde{x}(t)u^T(t)\} = G(t)Q(t)G^T(t) + K(t)R(t)K^T(t).$$

З урахуванням цього матричне диференціальне рівняння для коваріаційної матриці похибки матиме вигляд:

$$\begin{aligned} \dot{\tilde{P}}(t) = [F(t) - K(t)H(t)]\tilde{P}(t) + \tilde{P}(t)[F(t) - K(t)H(t)]^T + \\ + G(t)Q(t)G^T(t) + K(t)R(t)K^T(t). \end{aligned} \quad (6.4.8)$$

Так як $\hat{x}(t_0)$ є детермінованим вектором, рівним \bar{x}_0 , то в якості початкового значення матриці $\tilde{P}(t)$ приймають матрицю $P_0 = M\{x(t_0)x^T(t_0)\}$

Тепер можна перейти до мінімізації величини I , яка визначається співвідношенням (6.4.4), шляхом відповідного вибору матриці коефіцієнтів $K(t)$. Виявляється, що для розглядуваної задачі ця мета буде досягнута, якщо мінімізувати

$$\frac{dI(t)}{dt} = \text{tr}[\dot{\tilde{P}}(t)],$$

де $\tilde{P}(t)$ визначається рівнянням (6.4.8). При мінімізації цього скалярного показника шляхом підбору матриці $K(t)$ необхідно користуватись методами матричного числення.

Існує аналогія між операціями, які виконуються тут і у випадку, коли $K(t)$ і $\tilde{P}(t)$ є скалярними величинами. Маємо:

$$\begin{aligned}\frac{d\left(\operatorname{tr}\left\{\left[F(t)-K(t)H(t)\right]\tilde{P}(t)\right\}\right)}{dK(t)} &= -\operatorname{tr}\left\{H(t)\tilde{P}(t)\right\}, \\ \frac{d\left(\operatorname{tr}\left\{\tilde{P}(t)\left[F(t)-K(t)H(t)\right]^T\right\}\right)}{dK(t)} &= -\operatorname{tr}\left\{\tilde{P}(t)H^T(t)\right\}, \\ \frac{d\left(\operatorname{tr}\left\{G(t)Q(t)G^T(t)\right\}\right)}{dK(t)} &= 0, \\ \frac{d\left(\operatorname{tr}\left\{K(t)R(t)K^T(t)\right\}\right)}{dK(t)} &= \operatorname{tr}\left\{2K(t)R(t)\right\}.\end{aligned}$$

Враховуючи, що $\operatorname{tr}\{H(t)\tilde{P}(t)\} = \operatorname{tr}\{\tilde{P}(t)H^T(t)\}$, одержимо:

$$\frac{d\left(\operatorname{tr}\left[\dot{\tilde{P}}(t)\right]\right)}{dK(t)} = -2\operatorname{tr}\left\{\tilde{P}(t)H^T(t)\right\} + 2\operatorname{tr}\left\{K(t)R(t)\right\}.$$

Покладаючи

$$\frac{d\left(\operatorname{tr}\left[\dot{\tilde{P}}(t)\right]\right)}{dK(t)} = 0,$$

знайдемо вираз для матриці коефіцієнтів фільтру Калмана-Б'юсі —

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

Тоді для коваріаційної матриці мінімальної похибки одержимо наступне рівняння:

$$\begin{aligned}\dot{\tilde{P}}(t) &= F(t)\tilde{P}(t) + \tilde{P}(t)F^T(t) + G(t)Q(t)G^T(t) - \\ &- \tilde{P}(t)H^T(t)R^{-1}(t)H(t)\tilde{P}(t),\end{aligned}$$

яке називається **рівнянням Рікатті**.

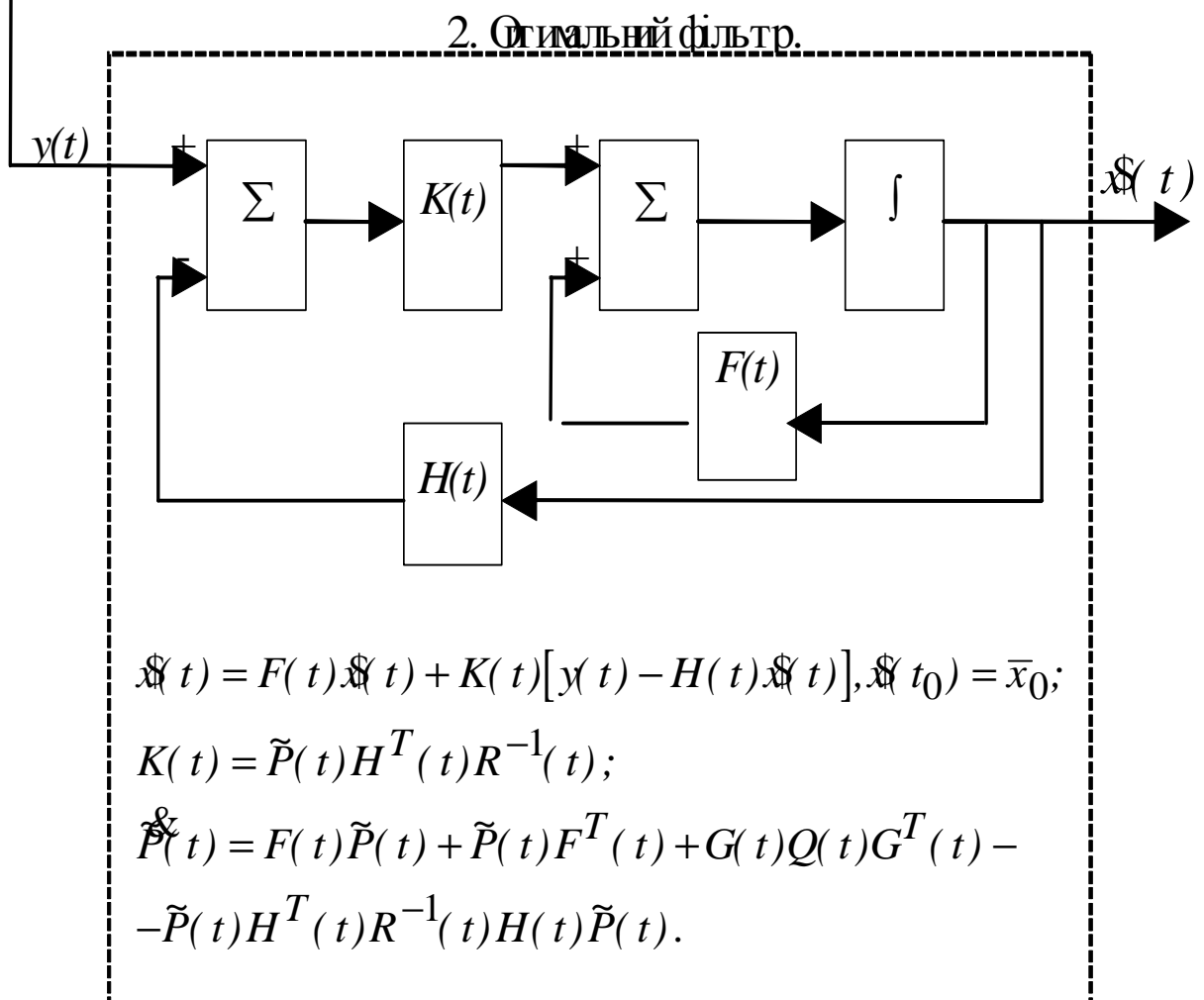
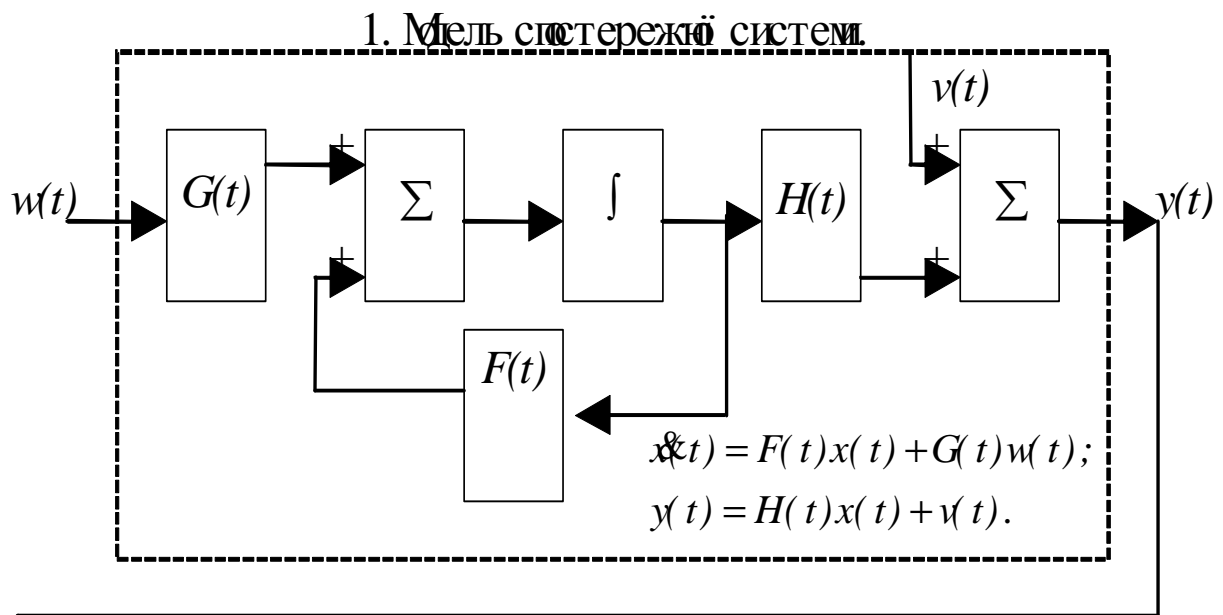
Цей результат завершує вивід рівнянь для оптимального лінійного фільтру, який формує незміщену оцінку з мінімальною дисперсією.

6.5. Структурна схема спостереження стану системи з мінімальною дисперсією похибки спостереження

Відповідна структурна схема наведена на мал. 6.1. Як бачимо, вона складається з двох частин:

- 1) моделі спостережної системи;
- 2) власне фільтру Калмана-Б'юсі.

Модель спостережної системи, тобто верхня обведена пунктиром область, закрита від нас, нам доступний лише сигнал $y(t)$ на виході цієї системи. Саме цей сигнал і поступає на вхід оптимального фільтру. Коефіцієнти $K(t)$ фільтру підібрані таким чином, щоб забезпечити мінімальну дисперсію похибки спостереження. На виході фільтру одержуємо оцінку вектора станів спостережної системи. На мал. 6.1. подані також необхідні співвідношення, які характеризують модель спостережної системи і оптимальний фільтр.



Мал. 6.1.

6.6. Алгоритм калманівської фільтрації

Процес оцінювання вектору станів за допомогою фільтру Калмана-Бьюсі зручно представити у вигляді алгоритму.

Алгоритм фільтру Калмана-Б'юсі:

1. Задаємо початкове значення оцінки вектора станів \hat{x}_0 і похибки оцінювання \tilde{x}_0 — так зване нульове наближення.

2. Розв'язуємо рівняння Рікатті:

$$\begin{aligned}\dot{\tilde{P}}(t) = & F(t)\tilde{P}(t) + \tilde{P}(t)F^T(t) + G(t)Q(t)G^T(t) - \\ & - \tilde{P}(t)H^T(t)R^{-1}(t)H(t)\tilde{P}(t),\end{aligned}$$

приймавши в якості початкового значення матрицю $\tilde{P}_0 = M \{\tilde{x}_0\tilde{x}_0^T\}$. В результаті одержимо матрицю $\tilde{P}(t)$.

3. Обчислюємо коефіцієнти фільтру:

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

4. Повертаємось на початок і шукаємо нове значення похибки $\tilde{x}(t)$.

Для цього спочатку знаходять

$$u(t) = G(t)w(t) - K(t)v(t),$$

а потім розв'язують рівняння:

$$\dot{\tilde{x}}(t) = [F(t) - K(t)H(t)]\tilde{x}(t) + u(t)$$

з початковою умовою \tilde{x}_0 .

5. Якщо похибка нас не задовільняє, процес обчислень проводимо спочатку, приймавши в якості \tilde{x}_0 нове значення похибки.

Якщо ж похибка нас задовільняє, знаходимо оцінку вектора станів, розв'язавши рівняння:

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + K(t)[y(t) - H(t)\hat{x}(t)]$$

з початковою умовою $\hat{x}(t_0) = \hat{x}_0$, де \hat{x}_0 — початкове значення оцінки вектора станів.

Отже, алгоритм калманівської фільтрації зводиться до того, що спочатку задається оцінка \hat{x}_0 з деякою, не обов'язково мінімальною, похибкою. Потім з допомогою ітеративної процедури ця оцінка покращується таким чином, щоб досягнути мінімально можливої середньоквадратичної похибки. На практиці обмежуються деяким заданим значенням похибки, при досягненні якого процес обчислень припиняють.

6.7. Особливості фільтру Калмана-Б'юсі

Розглядуваний фільтр має такі особливості:

- 1) апріорна інформація про сигнал (і шум) повинна задаватись в просторі станів;
- 2) алгоритм фільтрації має ітеративну форму і зручний для обробки цифровими обчислювальними машинами;
- 3) алгоритм фільтрації одночасно являє собою безпосередній опис способу реалізації фільтру;
- 4) легко поширюється на нестационарні сигнали; це стосується і випадку, коли спостереження починаються в довільний момент часу;
- 5) легко може бути поширений на багатомірний випадок.

У випадку стаціонарних процесів фільтр Калмана-Б'юсі переходить у фільтр Вінера.

Зауважимо, що зведення до рівняння Рікатті — один з можливих способів реалізацій фільтру Калмана.

Відомі інші методи — коли апріорно задаються розподіли. Тоді шукають оцінку станів за допомогою методу найбільшої правдоподібності.

Підсумок

У даному розділі розглянуто нестационарну систему (зі змінними параметрами) та ітераційну процедуру оптимальної оцінки її станів за апіорно заданими імовірнісними характеристиками. Ітераційна процедура полягає в покроковому наближенні з заданою точністю до значення відліку сигналу шляхом обчислень. Такі ланки (системи) є також екстремальними (оптимальними). Тут оптимальність забезпечується при оцінці параметрів ітераційної процедури. На відміну від Вінерівської фільтрації процедура є принципово чисельною (обчислювальною).

Розділ 7

АНАЛІЗ ОБРОБКИ ВИПАДКОВИХ СИГНАЛІВ НЕЛІНІЙНИМИ ЛАНКАМИ СИСТЕМИ

7.1. Статистична лінеаризація

Ідея лінеаризації використовується і у випадку стохастичних моделей сигналів при описові їх перетворень в нелінійних колах. Розглянемо лише кілька можливих методів аналізу нелінійних перетворень стохастичних сигналів, але таких, що стосуються різних моделей сигналів і стали "класичними":

а) перетворення випадкового сигналу в нелінійній неінерційній ланці кола, коли нелінійність задається неперервною, монотонною функцією, для якої існує обернена функція, а сигнал — нормально розподілений випадковий процес.

Нехай $x_1(t)$ — випадковий процес на вході нелінійної ланки, $x_2(t)$ — після нелінійного перетворення. Позначимо $F_{x_1}(x)$ функцію розподілу імовірностей значень сигналу на вході нелінійної ланки (густина розподілу імовірностей тоді позначається $f_{x_1}(x)$). Потрібно знайти функцію розподілу $F_{x_2}(x)$ імовірностей значень сигналу на виході нелінійної ланки (чи функцію густини $f_{x_2}(x)$).

При заданих умовах цієї задачі зауважимо наступне: якщо маємо функцію $Y = \varphi(x)$ (тут $Y \equiv Y(t)$, $X \equiv X(t)$) таку, що існує $\varphi^{-1}(Y) = X$, то для $F_1(x) = P(X < x)$ функція $F_2(y) = P(Y < y)$, при умові, що існує $f_1(x) = F_1'(x)$, буде рівною $\int_{\varphi(x) < y} f_1(x) dx$, бо

$$\frac{dF_2(y)}{dy} = \frac{dF_1(\varphi^{-1}(y))}{d\varphi^{-1}(y)} \left| \frac{d\varphi^{-1}(y)}{dy} \right|,$$
$$f_{x_2}(y) = f_{x_1}(\varphi^{-1}(y)) \left| \frac{d\varphi^{-1}(y)}{dy} \right|$$

(за відомим принципом інваріантності диференціала імовірності; тут $\left| \frac{d\varphi^{-1}(y)}{dy} \right|$ — Якобіан перетворення).

Тепер, якщо $Y = \varphi(x)$ — нелінійне перетворення, то при заданій функції розподілу імовірностей вхідного сигналу і деяких, вже названих інших умовах для математичного сподівання і дисперсії сигналу на виході нелінійної ланки маємо:

$$M[Y(t)] = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{-\infty}^{\infty} Y f_X(\varphi^{-1}(y)) \left| \frac{d\varphi^{-1}(y)}{dy} \right| dy$$

$$D[Y(t)] = \int_{-\infty}^{\infty} (y - m_y)^2 f_Y(y) dy = \int_{-\infty}^{\infty} (y - m_y)^2 f_Y(\varphi^{-1}(y)) \left| \frac{d\varphi^{-1}(y)}{dy} \right| dy.$$

В цих формулах опущено для спрощення позначення аргументу t . Розрахункові формули мають такий вигляд:

$$m_Y = \int_{-\infty}^{\infty} \varphi(x) f_X(x) dx;$$

$$\sigma_Y^2 = \int_{-\infty}^{\infty} (\varphi(x) - m_Y)^2 f_X(x) dx.$$

Зауважимо, що тут ідея лінеаризації "працює" на рівні диференціальних властивостей функцій розподілу. Нагадаймо, що сама ідея диференціального числення ґрунтується на властивостях лінійності;

б) обчислення коефіцієнтів лінеаризації. У цьому випадку сигнал на вході лінійної ланки представляється так:

$$X(t) = m_X + \overset{\circ}{X}(t),$$

де m_X — математичне сподівання, а $\overset{\circ}{X}(t)$ — центрований (випадковий) сигнал. Тоді сигнал на виході нелінійної ланки представляють у вигляді:

$$Y_1(t) = k_0 m_X + k_1 \overset{\circ}{X}(t),$$

де Y_1 — означає наближений до Y сигнал, k_0 — статистичний коефіцієнт підсилення математичного сподівання, k_1 — статистичний коефіцієнт підсилення випадкової складової вхідного сигналу.

Цей метод статистичної лінеаризації полягає в заміні нелінійної ланки еквівалентною (в імовірнісному сенсі) лінійною ланкою. Це означає, що:

$$MY_1(t) \cong MY(t); DY_1(t) \cong DY(t).$$

Тому маємо (аргумент t , де це можливо для спрощення записів, опускаємо):

$$MY_1 = M \left[k_0 m_X + k_1 \overset{\circ}{X} \right] = k_0 m_Y;$$

$$MY = \int_{-\infty}^{\infty} \varphi(x) f_X(x, t) dx;$$

$$DY_1 = M \left[k_1 \overset{\circ}{X} \right]^2 = k_1^2 \sigma_X^2;$$

$$DY = \int_{-\infty}^{\infty} (\varphi(x) - m_Y)^2 f_X(x, t) dx,$$

$$\text{звідки } k_0 = \frac{m_Y}{m_X}; (k_1)^2 = \frac{DY}{DX} = \frac{\sigma_Y^2}{\sigma_X^2};$$

за умови до похибки $M[Y - Y_1]^2 = \min$ маємо:

$$M[Y - Y_1]^2 = M(Y - k_0 m_X - k_1 \overset{\circ}{X})^2 = M[Y^2] + k_0^2 m_X^2 + k_1^2 M[(\overset{\circ}{X})^2] - \\ - 2k_0 m_X M[Y] - 2k_1 M\left[Y \overset{\circ}{X}\right] + 2k_0 k_1 m_X M\left[\overset{\circ}{X}\right].$$

Продиференціювавши цей вираз по k_0 і k_1 , прирівнявши до нуля (умова екстремуму), маємо, що:

$$\begin{cases} 2k_0^2 m_X - 2m_X m_Y = 0, \\ 2k_1 \sigma_X^2 - 2M\left[\overset{\circ}{X} Y\right] = 0. \end{cases}$$

Звідки отримаємо розрахункові формули:

$$k_0 = \frac{m_Y}{m_X}; \quad k_1^2 = \frac{M\left[Y \overset{\circ}{X}\right]}{\sigma_X^2} = \frac{M\left[\overset{\circ}{X} Y\right]}{DX}.$$

Отже, маємо два варіанти коефіцієнтів статистичної лінеаризації: (1) — коли вимагаємо рівності моментів при лінеаризації і (2) — коли вимагаємо мінімуму середньоквадратичної похибки статистичної лінеаризації. Причому, $k_0^{(1)} = k_0^{(2)}$. Для обчислення коефіцієнтів статистичної лінеаризації (задача (б)), необхідно розв'язати задачу (а).

7.2. Застосування функціональних рядів

для опису нелінійних ланок систем

Поряд з описом випадкових сигналів за допомогою функцій розподілу і її моментів є інші математичні моделі — інтегральні, які використовують ідею "чорного ящика", ідею зображень вигляду "вхід — вихід" у різних варіантах (лінійні випадкові процеси та їх ізоморфізми — вигляду перетворення Фур'є). Ця загальна ідея переноситься і на моделювання нелінійних ланок, зокрема, інерційних — моделі у вигляді функціонального ряду (ряду Вольтера).

Розглянемо однорідний стохастичний функціонал n -го порядку:

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_n(\tau_1, \dots, \tau_n; t) d\eta(\tau_1) \dots d\eta(\tau_n),$$

де K_n — ядро, дійсне, інтегровне з квадратом з вагою; $M|d\eta(\tau_1) \dots d\eta(\tau_n)|$ функція, M — математичне сподівання, η — випадковий процес з незалежними (ортогональними) приростами, відомий у подібних (одновимірних) зображеннях як порідний процес; так що у даному випадку маємо порідне поле.

Частинний випадок — $K_n(\cdot; t) = \sum_{k=1}^N \alpha_k \prod_{j=1}^n \varphi_{kj}(\tau_j, t)$, α_k — дійсні

невипадкові числа, а $\varphi_{kj}(\cdot, \cdot)$ — дійсна інтегровна з квадратом функція.

Розглядаючи поліном $\Phi_n(t)$, у який входять стохастичні функціонали різних порядків, приходимо до необхідності розглядати лінійно-незалежні $\Phi_n(t)$ при зображенні нелінійних перетворень стохастичних сигналів, представлені у вигляді лінійних інтегральних зображень, тобто виразів вигляду $M[\Phi_n(t)\Phi_m(t)]$. Крім того, для наближень необхідні нескінченні послідовності лінійно-незалежних $\Phi_n(t)$. Зауважимо, що нескінченну послідовність поліномів Φ_n можна ортогоналізувати процедурою Грама-Шмідта. Але, подібно як у випадку рядів типу Фур'є (при оптимізації у гільбертовому просторі за критерієм мінімальної норми різниці між зображенням у вигляді лінійної форми і підпростором) можна показати, що оптимальність наближення (апроксимації) досягається при властивості

ортогональності базисних елементів, тобто слід **вимагати** ортогональності поліномів Φ_n . А це призводить до введення спеціальної індикаторної функції:

$$I_n(\tau_1, K, \tau_n) = \begin{cases} 1, & \tau_1 \neq \tau_2 \neq K \neq \tau_n; \\ 0, & \text{вс юди.} \end{cases}$$

Зауважимо, що індикаторні функції означають випуклість множини, яка гарантує існування глобального оптимуму.

Отже, отримаємо:

$$G_n(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_n(\tau_1, \dots, \tau_n; t) I_n(\tau_1, \dots, \tau_n) \prod_{j=1}^n d\eta(\tau_j),$$

$G_0(t) = 1, n = \overline{1, \infty}$ — послідовність ортогональних стохастичних функціоналів.

$$\begin{aligned} M[G_n(t)G_m(t)] &= \delta_{nm} n! \chi_{\alpha}^n[\eta(I)] \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_n(\tau_1, \dots, \tau_n; t) \times \\ &\times \tilde{K}_n(\tau_1, \dots, \tau_n; s) \prod_{j=1}^n d\tau_j, \\ \delta_{nm} &= \begin{cases} 1, & n = m; \\ 0, & n \neq m, \end{cases} \end{aligned}$$

χ_{α}^n — семіінваріант, значок “ \sim ” — означає симетризацію. Норма функціоналу:

$$\begin{aligned} \|G_n(t)\| &= \left[n! \chi_{\alpha}^n[\eta(I)] \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_n^2(\tau_1, \dots, \tau_n; t) \prod_{j=1}^n d\tau_j \right]^{\frac{1}{2}}, \\ \overline{G_n(t)} &= G_n(t) [\|G_n(t)\|]^{-1}, \quad n = 0, 1, \dots \end{aligned}$$

Тоді відгук нелінійного пристрою (ланки) можна записати так:

$$y(t) = \sum_{n=0}^{\infty} C_n(t) \overline{G_n}(t),$$

де $C_n(t) = M[y(t)G_n(t)]$ — узагальнений коефіцієнт Фур'є розкладу $y(t)$. Звідси, для математичного сподівання і кореляційної функції відгуку $y(t)$ нелінійної ланки маємо:

$$\begin{aligned} m(t) &= My(t) = \sum_{n=0}^{\infty} C_n(t) M \overline{G_n}(t) = C_0(t), \\ B_y(t_1, t_2) &= M[y(t_1) - m(t_1)][y(t_2) - m(t_2)] = \\ &= \sum_{n=0}^{\infty} C_n(t_1) C_n(t_2) M[\overline{G_n}(t_1) \overline{G_n}(t_2)], \quad t \in T. \end{aligned}$$

Зауважимо, що для обчислення моментних функцій відгуку нелінійної ланки на випадкове збурення використано не функції розподілу, а не випадкові ядра, або послідовність семіінваріантів випадкового процесу.

Приклад. Нехай нелінійне перетворення має вигляд

$$F(x) = \sum_{n=1}^p a_n x^n,$$

$\{a_n, n = \overline{1, p}\}$ — дійсні числа; $\varphi(\tau, t) = e^{-\alpha(t-\tau)} u(t-\tau)$, $u(\cdot)$ — функція Гевісайда, $\eta(\tau) = \gamma(t)$ — гама-процес з незалежними приростами і послідовністю семіінваріантів $\chi_n[\gamma(\tau)] = \varepsilon^n \beta^n (n-1)! |\tau|$, $n = 2, 3, \dots$; $\chi_1[\gamma(\tau)] = 0$, β — параметр гама-процесу, $\varepsilon \equiv \text{sign}(\tau)$.

Знайти кореляційну функцію і математичне сподівання відгуку нелінійної ланки .

Вираз для відгуку:

$$y(t) = F[\xi(t)] = \sum_{n=1}^p a_n \xi^n(t) = \sum_{n=1}^p a_n \Phi_n(t),$$

де

$$\Phi_n(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{i=1}^n e^{-\alpha(t-t_i)} u(t-t_i) d\gamma(\tau_i).$$

Ортогональна система функціоналів

$$G_n = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{i=1}^n (e^{-\alpha(t-\tau_i)} u(t-\tau_i)) I_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n d\gamma(\tau_i),$$

$$n = 1, 2, \dots, G_0(t) = 1.$$

Норма для такої системи функціоналів —

$$\|G_n\| = \beta^n \left[\frac{n!}{(2\alpha)^n} \right]^{\frac{1}{2}}, n = 0, 1, 2, \dots$$

Ортонормована система виглядає так:

$$\overline{G}_n(t) = \frac{(2\alpha)^{\frac{n}{2}}}{\frac{1}{(n!)^{\frac{1}{2}} \beta^n}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{i=1}^n \left[e^{-\alpha(t-t_i)} u(t-t_i) \right] I_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n d\gamma(\tau_i),$$

$$n = 1, 2, \dots; \overline{G}_0(t) = 1.$$

Шуканий відгук у вигляді ортогонального ряду —

$$y(t) = \sum_{n=1}^p \alpha_n \sum_{m=0}^n C_{nm} \overline{G}_m(t),$$

коефіцієнти цього ряду —

$$C_{nm}(t) = M[\Phi_n(t)\overline{G_m(t)}] =$$

$$= C_{nm} = \begin{cases} \frac{(2\alpha)^{\frac{m}{2}}}{\beta^n} n!(m!)^{\frac{1}{2}} \sum_{\sum_i n_j p_j = n+m} \prod_{j=1}^v \frac{\beta_j^{n_j p_j}}{(p_j)^{2n_j} n_j!} \frac{1}{\alpha^n}, & m < n; \\ 0, & m > n. \end{cases}$$

Зауважимо, що вони не залежать від часу (стаціонарні), а також те, що доданки з $\chi_{pj}[\gamma(I)]$ при $p_j = 1$ рівні нулеві.

Отже, в результаті одержимо:

$$m(t) = m = \sum_{n=1}^p a_n C_{n0} = \sum a_n n! \sum_{\sum n_j p_j = n} \prod_{j=1}^v \frac{\beta_j^{n_j p_j}}{(p_j)^{2n_j} n_j!} \frac{1}{\alpha^{n_j}},$$

$$B_y(s) = \sum_{n_1, n_2}^p \alpha_{n_1} \alpha_{n_2} \sum_{m=1}^{\min(n_1, n_2)} C_{n_1 m} C_{n_2 m} e^{-m\alpha|s|}, \quad s = t_2 - t_1.$$

Перетворення Фур'є $B_y(s)$ дасть нам спектральну густину потужності:

$$S_y(\omega) = 4\alpha \sum_{n_1, n_2=1}^p \alpha_{n_1} \alpha_{n_2} \sum_{m=1}^{\min(n_1, n_2)} C_{n_1 m} C_{n_2 m} \frac{m}{m^2 \alpha^2 + \omega^2}, \quad \omega \in [0, \infty).$$

Щодо трудоемкості обчислень, то вона найбільша при обчисленні коефіцієнтів $C_n(t)$. І ще одне зауваження: аналіз впливу нелінійної ланки можна проводити за допомогою імітаційної моделі на ЕОМ, моделюючи і процес, і саму ланку. Але тоді необхідно проаналізувати затрати ресурсів (часу, пам'яті) і точності імітаційного моделювання. Досить часто застосовується комбінований підхід: аналітично (теоретично) отримують результат, що допомагає оцінити його вигляд, межі зміни величини, тощо, а сам аналіз проводиться на імітаційній моделі. Іноді кінцеві формули дають значний вииграш, але часом їх

чутливість до неточності в обчисленнях, представлення вхідних даних є ще значною. Отже, теоретичні дослідження є важливими, зокрема, і в даному конкретному випадку.

Підсумок

Вплив нелінійної ланки на стаціонарні випадкові процеси визначається подібно, до детермінованих випадків. Специфіку вносять зображення випадкового процесу, коли вимушені розглядати його характеристики (математичне сподівання, вищі моменти). Зауважимо, що термін "лінеаризація" тут використано в традиційному сенсі. Насправді, ніщо не "лінеаризується", а лише в найкращому (оптимальному, такому що забезпечує екстремальне значення критерію) вигляді нелінійна характеристика "підмінюється" лінійною. Вся подальша обробка проводиться як лінійна.

Застосування функціональних рядів дозволяє в асимптотичному сенсі, з використанням лінійних функціоналів вищих порядків, наблизитися до лінійних уявлень. При цьому ідея екстремальності використовується тут також: ортогоналізація системи функціоналів, симетрування їх ядер дозволяє в загальному зменшити кількість членів ряду, гарантує існування єдиного екстремуму — мінімального значення залишкового члену функціонального ряду.

Зауважимо, що у цьому розділі нелінійність вважається "корисною" характеристикою: проводиться аналіз впливу нелінійної ланки з константацією його факту. Оцінка впливу нелінійності як "шкідливого", як похибки, тут не розглядається.

В цілому, при виборі методу аналізу нелінійних ланок, систем використовують різноманітні "спрощення" в основі яких лежать конкретні властивості (моделі) процесу (синтезу) і нелінійності ланки (співвідношення між часами кореляції процесу і функції відгуку системи, величина інтенсивності випадкового впливу, тощо). Перелік і аналіз деяких таких випадків описано в книзі В. И. Тихонова "Нелинейные преобразования случайных процессов". Зокрема, коли інтенсивність випадкового процесу велика, а часи кореляції одного порядку, то це найбільш складний випадок і при таких умовах деякі нелінійні системи можна аналізувати за допомогою функціональних рядів Вольтерра. При великій інтенсивності випадкового сигналу і часові кореляції системи набагато більшому за час кореляції сигналу можна застосувати маркову модель (наприклад, аналіз динамічних систем проводити за допомогою рівняння ФПК). При випадкових

впливах малої інтенсивності добрі результати отримано методом лінійного наближення (лінеаризації).

ПІДСУМКИ

Радіоелектронна апаратура та системи, зокрема — ергатичні, коли у своїй діяльності їх використовує людина, яка трактується як ланка, проектується з заувагою моделі явища та сигналів, що супроводжують це явище. Інколи модель має описувати одночасно безліч факторів. Виявляється, що цього можна досягти, використавши імовірнісні методи, причому, немає сенсу зауважувати всі фактори. Щоб коректно застосувати імовірнісну модель при розробці радіо-електронної апаратури у цій частині посібника послідовно розглянуто її основні поняття: випадкові події (розділ 1), випадкової величини (розділ 2), випадкового процесу (розділ 3). Крім цього, відокремлено два підходи до опису випадкового процесу: а) за допомогою функції розподілу імовірностей, що логічно виникає з розгляду імовірності як такої (як характеристики випадкової події), "сім'ї" імовірностей (як параметризованої значеннями величини імовірностей або функції значень величин, тобто, власне функції розподілу імовірностей) і нарешті, "сім'ї" випадкових величин або ж функцій розподілу імовірностей (випадкового процесу); б) за допомогою спеціальних характеристик параметрів функції розподілу — моментів, як приклад, окремий випадок — теорії 2-го порядку.

Конкретні функції розподілу з'являються або за результатами наближень експериментальних даних, або (теоретично) — за результатами розв'язування оптимізаційних задач (за критерієм максимуму) виразу ентропії при обмеженнях на моменти функції розподілу та інтегралу від неї.

Задля використання сигналу маємо мати такі властивості його характеристик, що є або незмінними у часові, або їх зміни є детермінованими. Зокрема, загальною назвою таких властивостей може бути слово "стаціонарність".

Далі, у простих випадках виникає необхідність в інтерполяції, згладжуванні (апроксимації), прогнозові, виділенні сигналу з суміші (фільтрація) або спостережні неспостережних станів системи за спостережними станами. У імовірнісному випадкові це означає визначення відповідних функцій розподілу, кореляційних функцій, спектральних густин потужності або ж значень сигналу за вимірними значеннями сумішей. Для цього використовують апарат

функціонального аналізу, оптимізаційні методи, варіаційне числення, тощо. При цьому одержують або відразу сигнал (ітераційні методи), або характеристики обробки (фільтру, який ще потрібно реалізувати — тут вже справа за схемотехнікою).

Звичайно, вся різноманітність, всі комбінації можливих випадків описати у посібнику неможливо. Проте, описано ті, що стали вже "класичними".

Розглядають окремі задачі аналізу та синтезу. Проблеми реалізації в даному посібнику не описані.

КОНТРОЛЬНІ ПИТАННЯ

1. Імовірнісний простір.
2. Імовірнісна модель експерименту з зліченим числом результатів.
3. Суть теореми Каратеодорі про продовження міри. Загальна характеристика означень та теорем про борелівські класи множин.
4. Імовірнісна модель експерименту з нескінченним числом результатів.
5. Аксиоми теорії імовірностей.
6. Умовна імовірність. Формула повної імовірності. Формула Байєса.
7. Випадкова величина. Функція розподілу імовірностей випадкової величини.
8. Поняття про інтеграл. Інтеграли Лебега та Стільт'єса.
9. Математичне сподівання і дисперсія випадкової величини. Моменти. Моменти вищих порядків.
10. Характеристична функція.
11. Поняття про якість оцінювання параметрів функцій розподілу.
12. Опис випадкових процесів (сім'я випадкових величин).
13. Евклідов простір випадкових величин.
14. Кореляція. Геометрична інтерпретація кореляції.
15. Рівняння регресії для системи випадкових величин.
16. Поняття про збіжність послідовностей випадкових величин.
17. Моделі випадкового процесу.
18. Неперервність та диференційовність випадкових процесів. Інтегрування випадкових процесів.
19. Стаціонарний випадковий процес. Спектральна густина потужності стаціонарного процесу.
20. Приклади стаціонарних випадкових процесів.
21. Лінійні перетворення стаціонарних випадкових процесів. Спектральний розклад слабостаціонарного процесу.
22. Стаціонарні випадкові послідовності.
23. Загальна схема обробки сигналів у системах.
24. Закон розподілу відгуку лінійної системи на випадковий сигнал. Нормалізація випадкового процесу.
25. Енергетичний спектр відгуку лінійної системи на стаціонарний випадковий процес.

- 26.Кореляційна функція відгуку лінійної системи на стаціонарний випадковий процес.
- 27.Вплив на білий шум інтегруючої ланки.
- 28.Білий шум та вінерівський процес.
- 29.Визначення основних характеристик процесу на виході RC -ланки.
- 30.Визначення коефіцієнтів рівняння Фокера-Планка-Колмогорова.
- 31.Критерій оптимальності фільтру Вінера.
- 32.Блок-схема оптимального фільтру (фільтр Вінера).
- 33.Вивід рівняння Вінера-Хопфа методом варіаційного числення.
- 34.Вивід рівняння Вінера-Хопфа методом ортогоналізації похибки.
- 35.Розв'язування рівняння Вінера-Хопфа.
- 36.Похибка оптимальної фільтрації форми сигналу.
- 37.Задача Калманівської фільтрації.
- 38.Критерій оптимальності фільтру.
- 39.Поновлюючий процес у фільтрі Калмана-Б'юсі.
- 40.Коефіцієнти фільтру Калмана-Б'юсі.
- 41.Структурна схема спостереження стану системи з мінімальною дисперсією похибки спостереження.
- 42.Алгоритм калманівської фільтрації.
- 43.Особливості фільтру Калмана-Б'юсі.
- 44.Статистична лінеаризація.
- 45.Застосування функціональних рядів для опису нелінійних ланок систем.

ТИПОВІ ЗАДАЧІ

1. Обчислення імовірності події, яка може відбутися тільки разом з подією з повної групи подій (гіпотез).
2. Обчислення імовірності події з повної групи подій (гіпотез), коли мала місце подія, яка може відбутися тільки разом з подією — гіпотезою.
3. Обчислити імовірність появи події певну кількість разів у незалежних випробуваннях при відомій імовірності появи події в одному випробуванні.
4. Побудувати графік функції розподілу імовірності випадкової величини при заданих умовах чи результатах експерименту (для дискретної або неперервної випадкової величини).
5. Обчислити початкові та центральні моменти, числові характеристики випадкової величини при заданих функціях розподілу.
6. Знайти числові характеристики (математичне сподівання, дисперсію, коваріаційні моменти) системи випадкових величин.
7. Знайти числові характеристики функції випадкових величин при заданій функції та законах розподілу.
8. Знайти кореляційну функцію при заданому вигляді випадкового процесу (користуючись означенням кореляційної функції).
9. Знайти кореляційну функцію при заданій спектральній густині випадкового процесу.
10. Знайти спектральну густину при заданій кореляційній функції випадкового процесу.

11. Перевірити, чи заданий випадковий процес буде стаціонарним.

12. Знайти спектральну густину (кореляційну функцію) випадкового процесу на виході лінійної стаціонарної системи за кореляційною функцією (спектральною густиною) стаціонарного випадкового процесу на її вході:

- а) система задана у просторі змінних станів;
- б) задана функція передачі;
- в) задана імпульсна перехідна функція;
- г) заданий опис системи у вигляді диференціального (різницевого) рівняння.

13. Знайти функцію передачі (або імпульсну перехідну функцію) лінійної стаціонарної системи на вхід якої поступає адитивна суміш сигналу та завади (стаціонарні з нульовим математичним сподіванням, взаємно некорельовані випадкові процеси), таку, щоб на її виході спостерігати сигнал з мінімальною дисперсією похибки, коли:

- а) відомі спектральні густини сигналу та завади;
- б) задані відповідні автокореляційні функції;
- в) задані кореляційна функція сигналу і спектральна густина шуму.

14. Обчислити середньоквадратичну похибку оптимальної фільтрації при одній з умов, перелічених у п. 13, а-в.

15. Побудувати алгоритм фільтру Калмана.

16. Знайти оцінку з мінімальною дисперсією похибки вектора стану лінійної стаціонарної системи, заданої у просторі змінних стану і заданих кореляційних функціях вектора спостереження та сигналу на вході (система не вища другого порядку).

17. Побудувати фазовий портрет заданої нелінійної системи.

18. Обчислити коефіцієнти лінеаризації при заданій (описаній) нелінійності.

- 19.** Знайти сигнал на виході заданої нелінійної системи при заданому детермінованому сигналі на її вході, користуючись:
- а) методом лінеаризації;
 - б) рядами Вольтера.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

Балакришнан А. Введение в теорию оптимизации в гильбертовом пространстве. М.: Мир, 1974. —260 с.

Браммер К., Зиффлинг Г. Фильтр Калмана-Бьюси. М.: Наука, 1989. —200 с.

Гихман И. И., Скороход А. В., Ядренко М. И. Теория вероятностей и математическая статистика. Киев: Вища школа, 1979.—408 с.

Горячев В. Т., Журавлев А. Г., Тихонов В. И. Примеры и задачи по статической радиотехнике. М.: Сов. радио, 1970. —600 с.

Драган Я. П., Омельченко В. О. і ін. Прикладна теорія випадкових процесів і полів. Харків-Львів-Тернопіль, 1993.—248 с.

Казаков В. А. Введение в теорию марковских процессов и некоторые радиотехнические задачи. М.: Сов. радио, 1973.—232 с.

Марченко Б. Г. Методы стохастических интегральных представлений и его приложение в радиотехнике. Киев: Наукова думка, 1973.—192 с.

Марченко Б. Г., Омельченко В. А. Вероятностные модели случайных сигналов и полей в прикладной статистической радиофизике. Киев: УМК ВО, 1988.—176 с.

Марченко Б. Г., Щербак Л. Н. Линейные случайные процессы и их приложение. Киев: Наукова думка, 1975.—144 с.

Первозванский А. А. Случайные процессы в нелинейных автоматических системах. М.: Физматгиз, 1962.—352 с.

Пугачев В. С. Теория случайных функций и её применение к задачам автоматического управления. М.: ГИТТЛ, 1957.—660 с.

Ширяев А. Н. Вероятность. М.: Наука, 1980.—576 с.
