

Глава 1. Некоторые сведения из теории вероятностей и теории случайных процессов

§ 1.1. Вероятностные пространства и случайные величины

Определим некоторые математические понятия, соответствующие названию данного раздела [3, 19, 20, 27].

Математической моделью для случайного количества является *случайная величина*. Дадим ряд определений теории вероятностей.

Пусть задано пространство Ω элементов ω . В пространстве Ω выделим семейство множеств \mathfrak{F} , образующих σ -алгебру. Это означает, что для любого конечного или счетного набора множеств $A_i \in \mathfrak{F}$ их пересечение $\bigcap_i A_i$, сумма $\bigcup_i A_i$ и дополнения \bar{A}_i множеств A_i на Ω принадлежат \mathfrak{F} .

1.1.1. Определение. Пусть Ω - заданное множество, тогда σ -алгебра на Ω есть семейство \mathfrak{F} подмножеств множества Ω со следующими свойствами

- (i) $\emptyset \in \mathfrak{F}$;
- (ii) $F \in \mathfrak{F} \Rightarrow F^C \in \mathfrak{F}$, где $F^C = \Omega \setminus F$ - дополнение множества F в Ω ;
- (iii) $A_1, A_2, \dots \in \mathfrak{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathfrak{F}$.

Множества из \mathfrak{F} называются случайными событиями, а элементы $\omega \in \Omega$ - элементарными исходами.

Пара (Ω, \mathfrak{F}) называется измеримым пространством. Вероятностной мерой на измеримом пространстве (Ω, \mathfrak{F}) называется функция $P: \mathfrak{F} \rightarrow [0, 1]$, такая что

a) $P(\emptyset) = 0, P(\Omega) = 1$;

(b) если $A_1, A_2, \dots \in \mathfrak{F}$ и $\{A_i\}_{i=0}^{\infty}$ - непересекающаяся системами (т.е. $A_i \cap A_j = \emptyset$ при $i \neq j$), что

$$P\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Тройка $(\Omega, \mathfrak{F}, P)$ называется *основным вероятностным пространством*.
Вероятностным пространством называется полным, если \mathfrak{F} содержит все подмножества G множества Ω с P – внешней мерой нуль т.е. такие подмножества, что

$$P(G) = \inf \{ P(F); F \in \mathfrak{F}, G \subset F \} = 0.$$

Каждое вероятностное пространство может быть сделано полным добавлением к \mathfrak{F} всех множеств внешней меры нуль и продолжением меры P соответствующим образом.

Подмножества F множества Ω , которые принадлежат семейству \mathfrak{F} , называются \mathfrak{F} -измеримыми множествами. В вероятностной терминологии это означает, что

$$P(F) = \text{вероятность того, что произойдет событие } F.$$

В частности, если $P(F) = 1$, то это означает, что F произойдет с вероятностью 1 или «почти наверняка».

Теорема 1.1.1. Пусть на множестве \mathfrak{R} задана некоторая система его подмножеств D . Тогда существует наименьшая σ -алгебра, обозначаемая $\sigma(D)$, содержащая все множества из D .

Система $\sigma(D)$ является наименьшей в том смысле, что если \mathcal{A} – любая σ -алгебра подмножеств множества \mathfrak{R} , содержащая систему D , то $\sigma(D) \subseteq \mathcal{A}$.

Например, если \mathfrak{R} есть набор всех открытых подмножеств топологического пространства Ω (в частности, $\Omega = R^n$), то $\mathcal{B} = \sigma(D)$ называется *борелевской σ -алгеброй* на Ω , а элементы $B \in \mathcal{B}$ называются *борелевскими множествами*. Алгебра \mathcal{B} содержит все открытые множества, все счетные объединения замкнутых множеств, все счетные пересечения таких счетных объединений и т.д.

Случайным вектором $\xi(\omega)$ называется \mathfrak{F} -измеримая почти всюду конечная функция $\xi(\omega)$ со значениями в (R^n, \mathcal{B}) , т.е. $\xi(\omega) : (\Omega, \mathfrak{F}) \rightarrow (R^n, \mathcal{B})$ где \mathcal{B} – борелевская σ алгебра борелевских множеств пространства R^n .

Случайному вектору $\xi(\omega)$ можно сопоставить соответствующую функцию распределения $F(x_1, \dots, x_n)$, равную вероятности событий $(\xi_1 < x_1, \dots, \xi_n < x_n)$, где ξ_i - компонента вектора ξ . Иначе говоря,

$$F(x_1, \dots, x_n) = P(\xi_1 < x_1, \dots, \xi_n < x_n). \quad (1.1)$$

По любому из аргументов x_1, \dots, x_n при фиксированных остальных аргументах функция $F(x_1, \dots, x_n)$ монотонно не убывает.

Очевидны свойства функции распределения:

- 1) $F(-\infty) = 0$,
- 2) $F(+\infty) = 1$,
- 3) $F(x)$ при возрастании любой компоненты не убывает,
- 4) $F(x)$ может быть разрывной.

В ряде случаев функцию $F(x_1, \dots, x_n)$, когда эта функция непрерывна и дифференцируема, ее можно характеризовать с помощью плотности вероятности. Плотностью распределения вероятностей называется такая измеримая по Лебегу неотрицательная функция $f(x_1, \dots, x_n)$, что

$$F(x_1, \dots, x_n) = P(\xi_1 < x_1, \dots, \xi_n < x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(y_1, \dots, y_n) dy_1, \dots, dy_n. \quad (1.2)$$

Из этого равенства вытекает, что

$$f(x_1, \dots, x_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F(x_1, \dots, x_n). \quad (1.3)$$

Ясно, что интеграл от плотности вероятности $f(x_1, \dots, x_n)$ по всему пространству R^n равен единице, т.е.

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1, \dots, dx_n = 1. \quad (1.4)$$

Определение 1.1.2. Математическим ожиданием (или средним) случайной величины $\xi(\omega)$, определенной на $(\Omega, \mathfrak{F}, P)$, называется число

$$m_\xi = M(\xi) = \int_{\Omega} \xi(\omega) P(d\omega). \quad (1.5)$$

Математическое ожидание определено, если интеграл Лебега в правой части равенства (1.5) существует. Таким образом, математическое ожидание случайной величины ξ есть интеграл Лебега от функции $\xi(\omega)$ на Ω по вероятностной мере P . Если $P_\xi(\cdot)$ – закон распределения случайной величины $\xi(\omega)$ на $\mathcal{B}(R^1)$, а $F_\xi(y)$ – соответствующая функция распределения, то $M[\xi]$ можно вычислить следующим образом:

$$M[\xi] = \int_{R^1} y P_\xi(dy) = \int_{-\infty}^{\infty} y F_\xi(y).$$

Существование интегралов вытекает из существования математического ожидания, и наоборот, из существования интегралов следует существование математического ожидания.

Случайная величина называется центрированной, если $M[\xi] = 0$.

Основные свойства математического ожидания вытекают из свойства интеграла Лебега.

1. Математические ожидания $M[\xi]$ и $M[|\xi|]$ существуют и конечны одновременно, причем $M[\xi] \leq M[|\xi|]$.
2. Если $M[\xi]$ существует, то для любой конечной константы λ

$$M[\lambda\xi] = \lambda M[\xi].$$
3. Если $M[\xi]$ и $M[\eta]$ существуют и конечны, то $M[\xi + \eta] = M[\xi] + M[\eta]$.
4. Если $\xi \leq \eta$, то $M[\xi] \leq M[\eta]$.
5. Если $\xi \geq 0$ (P -п.н.), $M[\xi] < \infty$ и $\varepsilon > 0$, то $P(\xi \geq \varepsilon) \leq M[\xi]/\varepsilon$.
6. Если $M[|\xi|^p] < \infty$ при $p > 0$, то выполняется *неравенство Маркова*

$$P(|\xi| \geq \varepsilon) \leq M[|\xi|^p] / \varepsilon^p.$$

Если $\xi = \{\xi_1, \dots, \xi_n\}^T$ n -мерный вектор, то его математическим ожиданием называется вектор $M[\xi] = \{M[\xi_1], \dots, M[\xi_n]\}^T$.

Определение 1.1.3. Комплексная функция $\Phi_\xi(\lambda)$, $\lambda \in R^n$,

$$\Phi_\xi(\lambda_1, \dots, \lambda_n) = M[\exp\{i\lambda^T \xi\}] = \int_{R^n} \exp\{i\lambda^T x\} dF_\xi(x),$$

называется *характеристической функцией* распределения $F_\xi(x)$.

Теорема 1.1.4. Характеристическая функция однозначно определяет функцию распределения, т.е. если случайная величина ξ и случайная величина η имеют одну характеристическую функцию $\Phi_\xi(\lambda) = \Phi_\eta(\lambda)$, $\lambda \in R^n$, то $F_\xi(x) = F_\eta(x)$, $x \in R^n$.

Использование аппарата характеристических функций позволяет исследовать зависимость случайных величин в силу следующего утверждения.

Теорема 1.1.5. Случайные величины $\{\xi_i, i = 1, \dots, n\}$ независимы тогда и только тогда, когда $\Phi_\xi(\lambda_1, \dots, \lambda_n) = \prod_{i=1}^n \Phi_{\xi_i}(\lambda_i)$, где $\Phi_\xi(\lambda_1, \dots, \lambda_n)$ – характеристическая функция случайного вектора $\xi = \{\xi_1, \dots, \xi_n\}^T$, а $\Phi_{\xi_i}(\lambda_i)$, $\lambda_i \in R$ – характеристическая функция случайных величин $\xi_i, i = 1, \dots, n$.

Определение 1.1.6. Дисперсией $D[\xi]$ случайной величины $\xi(\omega)$ называется число $D[\xi] = D_\xi = M[(\xi - m_\xi)^2] = \int_{-\infty}^{\infty} (y - m_\xi)^2 dF_\xi(y)$.

Из определения и свойств интеграла Лебега следует:

1. $D[\xi] > 0$;
2. $D[a\xi + b] = a^2 D[\xi]$, если $a, b = const$;
3. $D[\{\sum_{k=1}^n \xi_k\}] = \sum_{k=1}^n D[\xi_k]$, если $D[\xi_k] < \infty$ и случайные величины $\{\xi_k\}$ независимы в совокупности;
4. $D[\xi] = M[\xi^2] - (m_\xi)^2$;
5. если $M[\xi^2] < \infty$, то выполняется *неравенство Чебышева*
 $P(|\xi - m_\xi| \geq \varepsilon) \leq D[\xi] / \varepsilon^2$.

Определение 1.1.7. Ковариацией случайных величин ξ и η называется величина

$$\text{cov}[\xi, \eta] = M[(\xi - m_\xi)(\eta - m_\eta)],$$

где $\overline{(\cdot)}$ – знак комплексного сопряжения.

Если $\xi = \{\xi_1, \dots, \xi_n\}^T$ n -мерный вектор, то его ковариационной матрицей называется матрица $R_\xi = \text{cov}[\xi_i, \xi_j]$. Очевидно, что $R_\xi = M[\xi \overline{\xi}^T] - m_\xi \overline{m_\xi}^T$.

Перечислим основные свойства ковариации.

1. Если $M[\xi^2] < \infty$, $M[\eta^2] < \infty$, то ковариация случайных величин ξ и η существует и удовлетворяет неравенству Коши – Буняковского: $\{\text{cov}[\xi, \eta]\}^2 \leq D[\xi]D[\eta]$.
2. Если $M[\xi] = M[\eta] = 0$, то $\text{cov}[\xi, \eta] = \{\xi\eta\}$ – скалярное произведение случайных величин ξ , η .
3. Если ξ и η независимы, то $\text{cov}[\xi, \eta] = 0$.
4. $D[\xi] = \text{cov}[\xi, \xi]$.
5. $D[\xi + \eta] = D[\xi] + D[\eta] + 2\text{cov}[\xi, \eta]$.

Если $\text{cov}[\xi, \eta] = 0$, то случайные величины ξ и η называются некоррелированными или ортогональными, что обозначается $\xi \perp \eta$.

Условное математическое ожидание. Пусть задано вероятностное пространство $(\Omega, \mathfrak{F}, P)$, \mathcal{G} – некоторая σ -алгебра случайных событий, т.е. $\mathcal{G} \subseteq \mathfrak{F}$, а ξ – случайная величина, такая, что $M[|\xi|] < \infty$.

Определение 1.1.8. Условным математическим ожиданием случайной величины ξ относительно \mathcal{G} называется случайная величина, обозначаемая $M[\xi/\mathcal{G}]$ и удовлетворяющая следующим условиям:

1. $M[\xi/\mathcal{G}]$ является измеримой;
2. для любого множества $A \in \mathcal{G}$ выполняется равенство

$$\int_A \xi(\omega) P(d\omega) = \int_A M[\xi/\mathcal{G}](\omega) P(d\omega). \quad (1.6)$$

Указанные два условия определяют условное математическое ожидание $\xi(\omega)$ однозначно с вероятностью 1. Условное математическое ожидание всегда существует, если только $M[|\xi|] < \infty$.

Приведем основные свойства условного математического ожидания, вытекающие из определения 1.1.4.

1. Если $\xi = C = const$ (P -н.н.), то $M[\xi / \mathcal{G}] = C$ (P -н.н.).
2. Если $\xi \leq \eta$ (P -н.н.), то $M[\xi / \mathcal{G}] \leq M[\eta / \mathcal{G}]$ (P -н.н.).
3. $M[\xi / \mathcal{G}] \leq M[|\xi| / \mathcal{G}]$ (P -н.н.).
4. Если a, b – константы, а ξ, η – случайные величины, такие, что $M[|\xi|] < \infty, M[|\eta|] < \infty$, то $M\{a\xi + b\eta\} / \mathcal{G} = aM[\xi / \mathcal{G}] + bM[\eta / \mathcal{G}]$ (P -н.н.).
5. Пусть $\mathcal{G} = \{\emptyset, \Omega\}$ – тривиальная σ -алгебра, тогда $M[\xi / \mathcal{G}] = M[\xi]$ (P -н.н.).
6. Если случайная величина ξ измерима относительно \mathcal{G} , то $M[\xi / \mathcal{G}] = \xi$ (P -н.н.).
7. $M[M[\xi / \mathcal{G}]] = M[\xi]$ (P -н.н.).
8. Если $\mathcal{G}_1 \subseteq \mathcal{G}_2$, то $M[M[\xi / \mathcal{G}_2] / \mathcal{G}_1] = M[\xi / \mathcal{G}_1]$ (P -н.н.).
9. Если $\mathcal{G}_2 \subseteq \mathcal{G}_1$, то $M[M[\xi / \mathcal{G}_2] / \mathcal{G}_1] = M[\xi / \mathcal{G}_2]$ (P -н.н.).
10. Если случайная величина ξ не зависит от \mathcal{G} , то $M[\xi / \mathcal{G}] = \xi$ (P -н.н.).
11. Пусть η измерима относительно \mathcal{G} и $M[\xi\eta] < \infty$, тогда $M[\xi\eta / \mathcal{G}] = \eta M[\xi / \mathcal{G}]$ (P -н.н.).
12. *Неравенство Иенсена.* Пусть $g(x)$ – выпуклая вниз функция, такая, что $M[|g(\xi)|] < \infty$, тогда $g(M[\xi / \mathcal{G}]) \leq M[g(\xi) / \mathcal{G}]$ (P -н.н.).
13. Пусть η произвольная измерима относительно \mathcal{G} случайная величина. Если $M[\xi^2] < \infty, M[\eta^2] < \infty$, то $M\{[\xi - M[\eta / \mathcal{G}]]^2\} \leq M\{[\xi - \eta]^2\}$ (P -н.н.).

Рассмотрим случайный вектор ξ с компонентами $\xi_i, i = 1, 2, \dots, n$. Если каждая компонента вектора может принимать дискретное множество значений $\xi_i(j_i)$, где $j_i = 1, 2, \dots, m_i$, то ясно, что они принимают (m_1, m_2, \dots, m_n) значений. Для характеристики случайного вектора можно задать совместное распределение вероятностей $p(j_1, j_2, \dots, j_n)$. Здесь $p(j_1, j_2, \dots, j_n)$ – вероятность того, что случайная величина ξ_1 принимает j_1 значение и случайная величина ξ_2 принимает j_2 значение и, . . . , и случайная величина ξ_n принимает j_n

значение. Если интересуются только одной компонентой случайного вектора, например ξ_1 , то безусловное распределение вероятностей случайной величины ξ_1 определяется как

$$p(j_1) = \sum_{j_2=1}^{m_2} \cdots \sum_{j_n=1}^{m_n} p(j_1, j_2, \dots, j_n). \quad (1.7)$$

Здесь $p(j_i)$ - вероятность такого сложного события, в котором ξ_1 принимает свое j_1 значение, а случайные величины $\xi_2, \xi_3, \dots, \xi_n$ принимают любые возможные значения. Итак, в зависимости от исходного пространства элементарных событий значение случайной величины может соответствовать результату либо одного элементарного события, либо сложного событий. В общем случае

$$p(j_1, j_2, \dots, j_i) = \sum_{j_{i+1}=1}^{m_{i+1}} \cdots \sum_{j_n=1}^{m_n} p(j_1, j_2, \dots, j_n). \quad (1.8)$$

Для характеристики отдельных компонент вектора ξ достаточно знать их безусловное распределение вероятностей. Однако для полного описания ξ нужно задать функцию распределение $p(x_1, x_2, \dots, x_n)$.

Как и в скалярном случае, вектор ξ можно приближенно характеризовать с помощью моментов распределения $p(\xi_1, \xi_2, \dots, \xi_n)$.

Вектор средних значений ξ определяется формулой

$$m_\xi = \bar{\xi} = M[\xi] = \sum_{j_1}^{m_1} \cdots \sum_{j_n}^{m_n} \begin{bmatrix} \xi(j_1) \\ \cdot \\ \cdot \\ \cdot \\ \xi(j_n) \end{bmatrix} p(j_1, j_2, \dots, j_n). \quad (1.9)$$

Рассмотрим вторые моменты вектора ξ . В отличие от единственной дисперсии случайной скалярной величины случайному вектору соответствует n дисперсий и $n(n-1)/2$ величин, называемых смешанными моментами второго порядка. Дисперсионная матрица D_ξ определяется как

$$D_{\xi} = M[(\xi - \bar{\xi})(\xi - \bar{\xi})^T] = M \begin{bmatrix} (\xi_1 - \bar{\xi}_1)^2 & \dots & (\xi_1 - \bar{\xi}_1)(\xi_n - \bar{\xi}_n) \\ \cdot & \dots & \cdot \\ (\xi_n - \bar{\xi}_n)(\xi_1 - \bar{\xi}_1) & \dots & (\xi_n - \bar{\xi}_n)^2 \end{bmatrix} =$$

$$= \sum_{j_1=1}^{m_1} \dots \sum_{j_n=1}^{m_n} \begin{bmatrix} (\xi_1 - \bar{\xi}_1)^2 & \dots & (\xi_1 - \bar{\xi}_1)(\xi_n - \bar{\xi}_n) \\ \cdot & \dots & \cdot \\ (\xi_n - \bar{\xi}_n)(\xi_1 - \bar{\xi}_1) & \dots & (\xi_n - \bar{\xi}_n)^2 \end{bmatrix} p(j_1, j_2, \dots, j_n). \quad (1.10)$$

Здесь предполагается, что оператор математического ожидания M применяется к каждому элементу матрицы. Диагональные элементы матрицы D_{ξ} являются дисперсиями компонент вектора; элементы, стоящие вне диагонали, есть смешанные моменты второго порядка. Дисперсионная матрица D_{ξ} симметричная и положительно определенная, так что имеется только $n(n+1)/2$ различных элементов.

Смешанные моменты второго порядка матрицы $\text{cov} [\xi_i, \xi_j]$ характеризуют степень линейной зависимости отдельных элементов случайного вектора друг от друга. Если $\text{cov} [\xi_i, \xi_j] \neq 0$ при $i \neq j$, то говорят, что случайные величины ξ_i и ξ_j коррелированы. Можно сказать, что случайные величины $\xi_1, \xi_2, \dots, \xi_n$ зависимы, если знание распределений $p(\xi_1), p(\xi_2), \dots, p(\xi_n)$ полностью не определяет распределение $p(\xi_1, \xi_2, \dots, \xi_n)$. С другой стороны, если

$$p(\xi_1, \xi_2, \dots, \xi_n) = p(\xi_1)p(\xi_2)\dots p(\xi_n) \quad (1.11)$$

для всех возможных значений $\xi_1, \xi_2, \dots, \xi_n$, то случайные величины называются независимыми. Следует отметить, что попарная независимость $\xi_1, \xi_2, \dots, \xi_n$ не означает независимости. В то же время попарная независимость является достаточной для некоррелированности. Даже отсутствие корреляции не означает независимости.

Пусть два случайных вектора ξ и η с дискретными значениями зависимы. Условная вероятность того, что произойдет событие η при конкретном значении ξ , определяется соотношением

$$P(\eta / \xi) = \frac{P(\eta, \xi)}{P(\xi)} \text{ для } P(\xi) \neq 0. \quad (1.12)$$

Условное среднее и условный второй момент определяется формулами (1.5) и (1.6). Совместное распределение вероятностей просто заменяется на условное распределение вероятностей. Однако поскольку математическое ожидание и второй момент есть функция случайной величины, то они уже являются не постоянными. Так как $P(\eta / \xi) = P(\xi / \eta)P(\eta)$, то

$$P(\eta / \xi) = \frac{P(\xi / \eta)P(\eta)}{P(\xi)}. \quad (1.13)$$

Это формула Бейеса. В этой формуле $P(\eta)$ - априорная вероятность события η без знания значения события ξ , а $P(\eta / \xi)$ - апостериорная вероятность события η при заданном значении события ξ . Когда оба события независимы, то выражение (1.13) приводит к равенству $P(\eta / \xi) = P(\eta)$, означающему, что, зная ξ , нельзя предсказать появление события η .

Для непрерывно распределенных случайных величин функция $P(\xi_1, \xi_2, \dots, \xi_n)$ становится плотностью распределения вероятностей случайных величин $f(\xi_1, \xi_2, \dots, \xi_n)$, причем $f(\xi_1, \xi_2, \dots, \xi_n)d\xi_1 d\xi_2 \dots d\xi_n$ есть вероятность того, что случайный вектор ξ будет находиться в элементарном объеме $d\xi_1 d\xi_2 \dots d\xi_n$ с центром в точке $(\xi_1, \xi_2, \dots, \xi_n)$. Таким образом, имеем

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(\xi_1, \xi_2, \dots, \xi_n) d\xi_1 d\xi_2 \dots d\xi_n = 1. \quad (1.14)$$

Математическое ожидание вектора определяется как

$$\bar{\xi} = m_{\xi} = M[\xi] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \begin{bmatrix} \xi_1 \\ \cdot \\ \cdot \\ \cdot \\ \xi_n \end{bmatrix} f(\xi_1, \xi_2, \dots, \xi_n) d\xi_1 d\xi_2 \dots d\xi_n = \begin{bmatrix} \bar{\xi}_1 \\ \cdot \\ \cdot \\ \cdot \\ \bar{\xi}_n \end{bmatrix}. \quad (1.15)$$

Дисперсия и смешанные моменты второго порядка определяются как

$$D = M[(\xi - \bar{\xi})(\xi - \bar{\xi})^T] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \begin{bmatrix} (\xi_1 - \bar{\xi}_1)^2 & \dots & (\xi_1 - \bar{\xi}_1)(\xi_n - \bar{\xi}_n) \\ \cdot & \dots & \cdot \\ (\xi_n - \bar{\xi}_n)(\xi_1 - \bar{\xi}_1) & \dots & (\xi_n - \bar{\xi}_n)^2 \end{bmatrix} f(\xi_1, \xi_2, \dots, \xi_n) d(\xi_1) d(\xi_2) \dots d(\xi_n).$$

Приведем часто встречающиеся совместные распределения.

Равномерное распределение. Равномерное распределение, когда все возможные значения равновероятны, является простейшим распределением случайного скаляра. Если имеется N возможных значений ξ , а именно $\xi(1), \xi(2), \dots, \xi(N)$, то

$$p(j) = \frac{1}{N}, \quad j = 1, 2, \dots, N.$$

Очевидно, что

$$\sum_{j=1}^N p(j) = 1, \quad \bar{\xi} = M[\xi] = \frac{1}{N} \sum_{j=1}^N \xi(j), \quad M[(\xi(j) - \bar{\xi})^2] = \frac{1}{N} \sum_{j=1}^N (\xi(j) - \bar{\xi})^2. \quad (1.16)$$

Биномиальное распределение. Пусть P - вероятность появления определенного события в каждом опыте. Тогда $1 - P$ - вероятность непоявления этого события. Вероятность того, что в n последовательных опытах событие не произойдет ни в одном из опытов, равна $P_0 = (1 - P)^n$. Вероятность того, что в n последовательных опытах это событие произойдет точно один раз, есть $P_1 = nP(1 - P)^{n-1}$, поскольку это событие может произойти в каждом из n опытов, и вероятность того, что это случится в любом конкретном опыте, а не во всех остальных, равна $P(1 - P)^{n-1}$. Аналогично вероятность того, что событие произойдет ровно k раз в n опытах, равна

$$P_k = \frac{n!}{k!(n-k)!} P^k (1 - P)^{n-k}, \quad (1.17)$$

так как $n! / k!(n-k)!$ есть число способов, какими из n различных элементов можно выбрать k элементов. Непосредственным вычислением получают среднее и дисперсию числа появлений события в n опытах

$$M[k] = \sum_{k=0}^n k P_k = nP, \quad M[(k - nP)^2] = \sum_{k=0}^n (k - nP)^2 P_k = nP(1 - P). \quad (1.18)$$

Распределение Пуассона. Рассмотрим биномиальное распределение, когда n очень велико, а P очень мало, например пусть $P = \mu/n$, где μ есть среднее число появлений события в n попытках. Вероятность того, что событие произойдет ровно k раз в большом количестве опытов, когда в среднем оно происходит μ раз, равна

$$P_k = \frac{\mu^k}{k!} e^{-\mu}. \quad (1.19)$$

Гауссовское распределение. Говорят, что случайный вектор $\xi(\omega) \in R^n$ имеет невырожденное гауссовское (или нормальное) распределение вероятностей, если существует вектор $m \in R^n$ и симметричная положительно-определенная матрица D такие, что для плотности вероятности $f(\xi)$ вектора ξ справедливо выражение

$$f(\xi) = \left[(2\pi)^n (\det D) \right]^{-1/2} \exp \left[-\frac{1}{2} (\xi - m_\xi)^T D^{-1} (\xi - m_\xi) \right], \quad \xi = (\xi_1, \dots, \xi_n)^T. \quad (1.20)$$

Из равенства (1.20), вычисляя соответствующие интегралы (1.5) и (1.6), можно установить, что математическое ожидание $M[\xi]$ гауссовского вектора $\xi(\omega)$ равно m_ξ , а дисперсионная матрица $M[(\xi - m_\xi)(\xi - m_\xi)^T] = D_\xi$.

Определение 1.1.9. Случайный вектор $\xi \in R^n$ имеет n -мерное гауссовское распределение с параметрами (m_ξ, K_ξ) , если его характеристическая функция имеет вид

$$\Phi_\xi(\lambda) = \exp \left\{ i\lambda^T m_\xi - \frac{1}{2} \lambda^T K_\xi \lambda \right\}, \quad \lambda \in R^n,$$

где m_ξ – математическое ожидание, а K_ξ – корреляционная матрица.

Последовательность случайных величин

Приведем основные определения и теоремы, относящиеся к вопросу о сходимости последовательностей случайных величин.

Будем далее считать, что случайные величины $\{\xi_n, n=1,2,\dots\}$ заданы на одном и том же вероятностном пространстве $(\Omega, \mathfrak{F}, P)$.

Определение 1.1.10. Последовательностью случайных $\{\xi_n\}$ величин будем называть *сходящийся почти наверное (P н.н.)* к случайной величине ξ , если $P\{\omega \in \Omega: \lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega)\} = 1$.

Определение 1.1.11. Последовательностью $\{\xi_n\}$ случайных величин называется *сходящейся по вероятности* к случайной величине ξ , если для любого $\varepsilon > 0$ $\lim_{n \rightarrow \infty} P\{|\xi_n - \xi| > \varepsilon\} = 0$.

Определение 1.1.12. Последовательностью $\{\xi_n\}$ случайных величин называется *фундаментальной по вероятности*, если для любого $\varepsilon > 0$ $\lim_{n, m \rightarrow \infty} P\{|\xi_n - \xi_m| > \varepsilon\} = 0$.

Определение 1.1.13. Последовательностью $\{\xi_n\}$ случайных величин называется *сходящейся в среднем порядка $d > 0$* к случайной величине ξ , если $\lim_{n \rightarrow \infty} P\{|\xi_n - \xi|^d\} = 0$.

При $d = 2$ этот вид сходимости называется *сходимостью в среднем квадратическом* и обозначается $\xi = \text{l.i.m.}_{n \rightarrow \infty} \xi_n$.

Определение 1.1.14. Последовательностью $\{\xi_n\}$ случайных величин называется *сходящейся по распределению (или слабо сходящейся)* к случайной величине ξ , если для любой равномерно ограниченной непрерывной функции $f(x)$, $x \in R^1$, $\lim_{n \rightarrow \infty} M[f(\xi_n)] = M[f(\xi)]$.

Если через $F_n(x)$ обозначить функцию распределения $\{\xi_n\}$, а через $F(x)$ функцию распределения случайной величины ξ , то для слабой сходимости последовательности $\{\xi_n\}$ к ξ необходимо и достаточно чтобы $F_n(x) \rightarrow F(x)$ при $n \rightarrow \infty$ в каждой точке $x \in R^1$, в которой функция $F(x)$ непрерывна. При этом, если $F(x)$ непрерывная функция, то $\{F_n(x)\}$ сходится к $F(x)$ равномерно на R^1 , т.е. $\sup_{x \in R^1} |F_n(x) - F(x)| \rightarrow 0$. $n \rightarrow \infty$.

Сходимость по распределению является наиболее слабой сходимостью последовательности случайных величин.

Теорема 1.1.15. Если последовательность $\{\xi_n\}$ сходится к ξ (P.n.n.) и найдется случайная величина η , такая, что $|\xi_n| \leq \eta$ (P.n.n.) для всякого n и $M[\eta] < \infty$, то $M[|\xi_n|] < \infty$, $M[\xi_n] \rightarrow M[\xi]$ и $M[|\xi_n - \xi|] \rightarrow 0$ при $n \rightarrow \infty$.

Теорема 1.1.16. Для сходимости последовательности $\{\xi_n\}$ по вероятности необходимо и достаточно, чтобы последовательность была фундаментальной (P.n.n.), т.е. для любого $\varepsilon > 0$ $\lim_{n \rightarrow \infty} P\{\sup_{k \geq 0} |\xi_{n+k} - \xi_n| > \varepsilon\} = 0$.

Приведенные результаты вытекают из теорем о предельном переходе под знаком интеграла Лебега и задают правила предельного перехода под знаком математического ожидания.

§ 1.2. Построение случайных процессов

Определение 1.2.1. Случайный процесс есть семейство (действительных или комплексных) случайных величин $\{\xi(t, \omega), t \in T\}$, определенных на $(\Omega, \mathfrak{F}, P)$ [15, 20, 21].

Множеством T обычно полупрямая $(0, \infty)$, однако это может быть и отрезок $[a, b]$, множество неотрицательных целых чисел и даже подмножества пространства R^n для $n \geq 1$. Отметим, что для каждого фиксированного $t \in T$ получаем случайную величину $\omega \rightarrow \xi_t(\omega)$, $\omega \in \Omega$. С другой стороны, фиксируя $\omega \in \Omega$, получаем функцию $t \rightarrow \xi_t(\omega)$, $t \in T$, которая называется траекторией процесса ξ_t . В приложениях непрерывный параметр t часто является мерой времени или расстояния.

Пусть B_1 есть σ -алгебра борелевских множеств отрезка $[t_0, T]$ такого, что $t \in [t_0, T]$ и $\omega \in U$. Обозначим через $B_1 \times U$ минимальную σ -алгебру множеств $T \times \Omega$, содержащую все подмножества вида $\Delta \times A$, где $\Delta \in [t_0, T]$, $A \in U$.

Функция $\xi(t, \omega)$ двух аргументов $t \in [t_0, T]$ и $\omega \in U$ со значениями в (R^n, B_1) называется *измеримой*, если она $(B_1 \times U)$ -измерима и при любом фиксированном $t \in [t_0, T]$ вектор $\xi(t, \omega)$ представляет собой случайный вектор.

Функция $\xi(t, \omega)$ при фиксированном $\omega \in U$ называется *траекторией* (или *реализацией* или *выборочной функцией*).

Обычно рассматривают сепарабельные случайные функции. Случайная функция называется сепарабельной, если ее поведение при всех $t \in [t_0, T]$ с точностью до событий вероятности нуль определяется поведением на каком-либо счетном всюду плотном множестве $\Delta \in [t_0, T]$. Известно, что для любой стохастически непрерывной (т.е. непрерывной в смысле сходимости по вероятности) функции $\xi(t, \omega)$ всюду на $[t_0, T]$, кроме может быть счетного множества точек, существует сепарабельная функция $\xi(t, \omega)$ такая, что $P\{\xi(t, \omega) = \xi_1(t, \omega)\} = 1, \forall t \in [t_0, T]$.

Для описания функции $\xi(t, \omega)$ можно использовать ее всевозможные конечномерные распределения. Обозначим через $A_i \in R^n$ произвольные борелевские множества, а через t_i ($i = 1, \dots, m$) - произвольные точки из интервала $[t_0, T]$.

Конечномерными распределениями вероятностей функции $\xi(t, \omega)$ называются вероятности

$$F(t_1, \dots, t_m, A_1, \dots, A_m) = P[\xi(t_1, \omega) \in A_1, \dots, \xi(t_m, \omega) \in A_m]. \quad (1.21)$$

Эти распределения удовлетворяют следующим двум очевидным свойствам:

1. Для любой перестановки π набора $(1, 2, \dots, m)$

$$F(t_{\pi(1)}, \dots, t_{\pi(m)}, A_{\pi(1)}, \dots, A_{\pi(m)}) = F(t_1, \dots, t_m, A_1, \dots, A_m).$$

Это просто означает, например, что $F(t_1, t_2, A_1, A_2) = F(t_2, t_1, A_2, A_1)$, так как представляет собой разные способы записи вероятности $P[\xi(t_1, \omega) \in A_1, \xi(t_2, \omega) \in A_2]$.

2. $\lim_{A_i} F(t_1, \dots, t_m, A_1, \dots, A_m) = F(t_1, \dots, t_{i-1}, \dots, t_{i+1}, \dots, t_m, A_1, \dots, A_{i-1}, \dots, A_{i+1}, \dots, A_m)$.

Свойства (1) и (2) известны как условия согласованности Колмогорова. Эти свойства достаточны для того, чтобы семейство функций распределения соответствовало некоторому случайному процессу.

1.2.2. Теорема Колмогорова. Пусть для каждого m и любых $t_1, \dots, t_m \in T$ задана функция распределения $F(t_1, \dots, t_m, A_1, \dots, A_m)$ и семейство этих функций удовлетворяет условиям согласованности (1) и (2). Тогда существуют вероятностное пространство $(\Omega, \mathfrak{F}, P)$ и случайный процесс $\xi(t, \omega)$ такие, что справедлива формула (1.21), т.е. функция $F(t_1, \dots, t_m, A_1, \dots, A_m)$ является семейством конечномерных распределений процесса $\xi(t, \omega)$.

Другими словами, в соответствии с теоремой Колмогорова для любого согласованного набора функций $F(t_1, \dots, t_m, A_1, \dots, A_m)$ существует такой случайный процесс, для которого эти функции служат конечномерными распределениями. Оказывается, что в общем случае такой случайный процесс не будет единственным, что означает, что семейство конечномерных распределений задает целый класс случайных процессов, которые в некотором смысле являются эквивалентными.

Пусть $\xi(t, \omega)$ и $\eta(t, \omega)$ - случайные процессы на $(\Omega, \mathfrak{F}, P)$. Говорят, что $\xi(t, \omega)$ есть версия (модификация) процесса $\eta(t, \omega)$, если $P\{\xi(t, \omega) = \eta(t, \omega)\} = 1, \forall t \in [t_0, T]$.

Отметим, что если процесс $\xi(t, \omega)$ есть версия процесса $\eta(t, \omega)$, то $\xi(t, \omega)$ и $\eta(t, \omega)$ имеют одинаковые конечномерные распределения. В этом смысле два таких процесса одинаковы, хотя свойства их траекторий могут быть разными.

Математическое ожидание случайной функции $\xi(t, \omega)$ определяется формулой

$$m_\xi = \bar{\xi}(t) = M[\xi(t, \omega)] = \int_{\Omega} \xi(t, \omega) P(d\omega).$$

Математическое ожидание есть среднее значение процесса.

«Центрированный» процесс $\xi(t, \omega) - \bar{\xi}(t)$ имеет нулевое среднее и так как он отличается от $\xi(t, \omega)$ только на детерминированную функцию, то часто удобнее рассматривать центрированный вариант процесса, а не сам процесс.

Предположим, что $M[\xi(t, \omega) \xi^T(t, \omega)]$ положительно определенная ограниченная матрица для всех t , т.е. $M\|\xi(t, \omega)\|^2 < \infty$. Отсюда следует, что $M\|\xi(t, \omega)\| < \infty$, так как если $F(\xi)$ - функция распределения $\xi(t, \omega)$, то

$$\begin{aligned} M\|\xi(t, \omega)\| &= \int_{-\infty}^{\infty} \|\xi(t, \omega)\| dF(\xi) = \\ &= \int_{\|\xi\| \leq 1} \|\xi(t, \omega)\| dF(\xi) + \int_{\|\xi\| > 1} \|\xi(t, \omega)\| dF(\xi) \leq \\ &\leq \int_{\|\xi\| \leq 1} dF(\xi) + \int_{\|\xi\| > 1} \|\xi(t, \omega)\|^2 dF(\xi) = 1 + M\|\xi(t, \omega)\|^2. \end{aligned}$$

Используя неравенство Шварца (или Коши - Буняковского), получим

$$M|\xi(t, \omega) \xi^T(s, \omega)| \leq \left\{ M\|\xi(t, \omega)\|^2 M\|\xi(s, \omega)\|^2 \right\}^{\frac{1}{2}}.$$

Так как $\xi(t, \omega)$ и $\xi(s, \omega)$ интегрируемы, то можно определить *матрицу корреляционных функций*

$$K_{\xi}(t, s) = M[\{\xi(t, \omega) - \bar{\xi}\}\{\xi(s, \omega) - \bar{\xi}\}^T].$$

Она является симметричной, определенной на $T \times T$.

Случайная функция называют гауссовской, если все ее конечномерные распределения вероятностей являются гауссовскими. Эти распределения зависят только от математического ожидания и корреляционной матрицы.

1.2.3. Предложение. Пусть $K(t, s)$, $t, s \in T$ - произвольная неотрицательно определенная функция. Тогда существует гауссовский процесс $\xi(t)$, $t \in T$ такой, что

$$K(t, s) = \text{cov}[\xi(t), \xi(s)].$$

Доказательство. Выберем произвольные $t_1, t_2, \dots, t_n \in T$. Матрица

$$Q = [K(t_i, t_j)] = \begin{bmatrix} K(t_1, t_1) & K(t_1, t_2) & \dots & K(t_1, t_n) \\ K(t_2, t_1) & K(t_2, t_2) & \dots & K(t_2, t_n) \\ \dots & \dots & \dots & \dots \\ K(t_n, t_1) & K(t_n, t_2) & \dots & K(t_n, t_n) \end{bmatrix}$$

является неотрицательно определенной по предположению, и поэтому характеристическая функция процесса $\xi(t)$

$$\Phi_{\xi}(\lambda_1, \dots, \lambda_n) = \exp\left\{-\frac{1}{2}\lambda^T Q \lambda\right\} \quad (1.22)$$

есть характеристическая функция гауссовского случайного вектора $\xi(t_1), \xi(t_2), \dots, \xi(t_n)$ с нулевым средним и ковариационной матрицей Q (Определение 1.1.9). Нетрудно проверить, что семейство конечномерных распределений, соответствующих характеристической функции (1.22), удовлетворяет условиям согласованности (i) и (ii). Следовательно, согласно теореме 1.2.2, существует процесс $\xi(t), t \in T$ с характеристической функцией (1.22), однозначно определяющей функцию распределения (теорема 1.1.4). Таким образом, $\xi(t)$ есть гауссовский процесс, а из его двумерных распределений получаем $\text{cov}[\xi(t), \xi(s)] = K(t, s)$.

Этот результат показывает, во первых, что существует взаимно однозначное соответствие между неотрицательно определенными функциями и гауссовскими процессами, а во вторых, что для любого процесса с конечным вторым моментом существует гауссовский процесс с той же ковариационной функцией. Это соответствует тому, что в конечномерном случае для любого случайного процесса $\eta(t)$ существует гауссовский случайный процесс $\xi(t)$ с той же корреляционной матрицей.

Полное описание случайного процесса общего вида потребовало бы знания всех возможных плотностей совместных распределений $p[\xi(t_1), \xi(t_2), \dots, \xi(t_N)]$ для всех $t_1, t_2, \dots, T \in [t_0, T]$, где N есть натуральное число.

Однако существуют случайные процессы, для которых какое-нибудь из многочисленных распределений определяет всю последовательность ее многомерных распределений. Примером случайных процессов, у которых все многомерные распределения определяются одним из них, являются марковские случайные процессы.

Определение 1.2.4. *Марковским случайным процессом (процессом Маркова)* называется вещественный случайный процесс $\{\xi_t\}_{t \in T}$, если для

любых $t_j \in T: t_1 < \dots < t_{k-1} < t_k$, $j=1, \dots, k$, произвольного целого $k > 1$ и любого борелевского множества $B \in \mathcal{B}(R^1)$ выполнено равенство

$$P\{\xi(t_k) \in B \mid \xi(t_1), \dots, \xi(t_{k-1})\} = P\{\xi(t_k) \in B \mid \xi(t_{k-1})\} \quad (P\text{-п.н.}). \quad (1.23)$$

Свойство (1.23) называется *марковским свойством*.

Соотношение (1.23) можно записать также в виде

$$P\{\xi(t_k) \in B \mid \xi(t_1) = x_1, \dots, \xi(t_{k-1}) = x_{k-1}\} = P\{\xi(t_k) \in B \mid \xi(t_{k-1}) = x_{k-1}\},$$

где $x_j \in R^1$ – произвольное допустимое значение случайной величины

$$\xi(t_j), \quad j=1, \dots, k.$$

Таким образом, вероятностное распределение состояния процесса в момент t_k зависит лишь от того, в каком состоянии находится процесс в ближайшем прошлом, т.е. при $t = t_{k-1}$, но не зависит от его состояний, предшествующих моменту времени t_{k-1} .

Определение 1.2.5. *Переходная вероятность марковского процесса* определяется как

$$P(s, x, t, B) = P\{\xi(t) \in B \mid \xi(s) = x\}$$

при $t > s$, $x \in R^1$, $B \in \mathcal{B}(R^1)$ и удовлетворяет соотношению

$$P(s, x, t, B) = \int_{R^1} P(s, x, u, dy) P(u, y, t, B),$$

которое выполняется для всех $s, u, t \in T: s \leq u \leq t$ и называется *уравнением Колмогорова-Чепмена*.

Функция $p(s, x, t, y)$ называется *переходной плотностью* распределения марковского процесса и при всех $s, t \in T$ и $x, y \in R^1$ удовлетворяет соотношениям

1. $p(s, x, t, y) \geq 0$ (условие неотрицательности);
2. $\int_{R^1} p(s, x, t, y) dy = 1$ (условие нормировки).

Случайные функции с независимыми приращениями

Функция $\xi(t, \omega)$, $t \in [t_0, T]$ называется функцией с *независимыми приращениями*, если случайные величины $\xi(t_1, \omega) - \xi(t_0, \omega), \dots, \xi(t_m, \omega) - \xi(t_{m-1}, \omega)$

взаимно независимы для любых $t_0 < t_1 < t_2 < \dots < t_{m-1} < t_m \leq T$. Любой процесс с независимыми приращениями представляет собой марковский случайный процесс.

В 1828 г. Шотландский ботаник Роберт Броун обнаружил, что взвешенные в жидкости частицы пыльцы совершают нерегулярные движения. Позднее эти движения были объяснены случайными столкновениями с молекулами жидкости. Важность броуновского движения для теории систем состоит в том, что оно дает хорошую математическую модель для определения типа шума, встречающегося в различных областях практики.

Гауссовский случайный процесс с независимыми приращениями и непрерывными реализациями называется *процессом броуновского движения* (или *винеровским процессом*). Реализации этого процесса являются непрерывными, нигде не дифференцируемыми функциями.

Винеровский процесс $\xi(t, \omega)$ называется стандартным, если $\xi(0, \omega) = 0$, $M[\xi(t, \omega)] = 0$, $M[\xi(t, \omega) \xi(t, \omega)^T] = I t$, где I - единичная матрица. Для винеровского процесса $\xi(t, \omega)$ с вероятностью 1 справедливо соотношение, называемое *законом повторного логарифма*:

$$\lim_{t \rightarrow \infty} \xi(t, \omega) (2t \ln \ln t)^{-1/2} = 1, \quad (1.24)$$

где \lim означает верхний предел.

Кроме того, траектории винеровского процесса имеют неограниченную вариацию на любом отрезке $[0, T]$, $T > 0$. Иными словами, для любого положительного числа δ справедливо соотношение

$$P \left[\sum_{j=0}^{n-1} |\xi(t_{j+1}, \omega) - \xi(t_j, \omega)| > \delta \right] \rightarrow 1$$

при $(t_{j+1} - t_j) \rightarrow 0$, $0 = t_0 < t_1 < \dots < t_n = T$.

§ 1.3. Пространство квадратично интегрируемых случайных величин

Пусть $(\Omega, \mathfrak{F}, P)$ - вероятностное пространство. Случайная величина X на $(\Omega, \mathfrak{F}, P)$ квадратично интегрируема, если $M[X^2] < \infty$. Множество H всех таких

случайных величин после некоторой модификации будет гильбертовым пространством. Очевидным образом определяются алгебраические операции, а скалярное произведение равно $\langle X, Y \rangle = M[XY]$.

Отсюда следует, что расстояние между X и Y есть

$$d\{X, Y\} = \|X - Y\| = \left\{ M[(X - Y)^2] \right\}^{\frac{1}{2}}.$$

Последовательность X_n сходится к X , если $d\langle X_n, X \rangle \rightarrow 0$, т.е., если $M[(X_n - X)^2] \rightarrow 0$. Такая сходимость называется сходимостью в среднем квадратическом. Следует отметить, что согласно аксиомам гильбертова пространства, если $d\langle X, Y \rangle = 0$, то $X = Y$. Однако $M[(X - Y)^2] = 0$ означает только, что $X = Y_{п.н.}$ (почти наверное), так, что X и Y могут различаться на множестве D таком, что $P(D) = 0$. Эта трудность является технической, которую можно обойти, считая такие величины идентичными. Формально рассматривается не само пространство \tilde{H} , а пространство \mathcal{H} классов эквивалентных случайных величин в \tilde{H} (две случайные величины X и Y эквивалентны, если $P(X = Y) = 1$). Заметим, что в силу предположения X^1, X и Y^1, Y суть две пары эквивалентных случайных величин, то $\langle X, Y \rangle = \langle X^1, Y^1 \rangle$, так что скалярное произведение определено корректно для элементов из \mathcal{H} . Пространство \mathcal{H} часто обозначается $L_2(\Omega, \mathfrak{F}, P)$ или просто L_2 , если существующее вероятностное пространство не вызывает сомнений. В общем случае множество случайных величин с $M|X|^p < \infty$ обозначается L_p . Пространство \mathcal{H} есть гильбертово пространство.

Введем ряд определений.

Определение 1.3.1. Непустое множество H называется линейным пространством, если на H определены две операции: сложение элементов и умножение элементов на число. Если умножение определено на вещественные числа, то пространство называется *вещественным* или *действительным*, если умножение определено на комплексные числа, то пространство называется *комплексным*.

Определение 1.3.2. Подпространством M линейного пространства H называется подмножество, замкнутое относительно операций сложения и умножения, т.е. $M \subset H$.

Определение 1.3.3. Линейной оболочкой $\mathcal{L}(M)$ некоторого подпространства $M \subseteq H$ называется множество всех возможных линейных комбинаций элементов множества M .

Определение 1.3.4. Пусть на H определено скалярное произведение. Нормой элемента $x \in H$ называется число $\|x\| = (x, x)^{1/2}$.

Определение 1.3.5. Линейное пространство с нормой (1.3.4) называется *полным*, если любая фундаментальная последовательность его элементов сходится к некоторому элементу этого пространства.

Определение 1.3.6. Линейное пространство \mathcal{H} со скалярным произведением и нормой (1.26) называется *гильбертовым пространством*, если оно полное.

Пространство \mathcal{H} часто обозначается $L_2(\Omega, \mathfrak{F}, P)$ или просто L_2 , если существующее вероятностное пространство не вызывает сомнений. В общем случае множество случайных величин с $M|X|^p < \infty$ обозначается L_p .

1.3.1. Оператор проектирования в гильбертовом пространстве

1.3.1. Теорема. Пусть \mathcal{H} - гильбертово пространство и $M \subset \mathcal{H}$ - его замкнутое подпространство. Тогда любой элемент $x \in \mathcal{H}$ имеет единственное разложение вида $x = y + z$,

где $y \in M, z \perp M$. Кроме того,

$$\|x - y\| = \min_{v \in M} \|x - v\|.$$

Доказательство. Чтобы доказать единственность, допустим, что существуют две пары элементов x, y и x_1, y_1 с требуемыми свойствами. Тогда $x = y + z = y_1 + z_1$, так что $y - y_1 = z_1 - z$. Но $y - y_1 \in M$, а $z_1 - z \in M^\perp$. Следовательно, $(y - y_1, y - y_1) = (y - y_1, z_1 - z) = 0$. Поэтому $\|y - y_1\| = 0$ и $y = y_1$ и $z = z_1$. Таким образом, самое большее одна пара x, y может удовлетворять

необходимым требованиям. Если $x \in M$, то $x = x + 0$ и есть искомое разложение. Поэтому пусть $x \notin M$. Так как M замкнуто, то в этом случае

$$\inf_{v \in M} \|x - v\| = h > 0. \quad (1.25)$$

Пусть Y_n - последовательность в M такая, что $\|y_n - x\| \downarrow h$. В силу тождества параллелограмма для $u, v \in \mathcal{H}$

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Полагая $u = y_n - x$, $v = y_m - x$, получаем

$$\|y_n + y_m - 2x\|^2 + \|y_n - y_m\|^2 = 2\|y_n - x\|^2 + 2\|y_m - x\|^2, \quad (1.26)$$

откуда следует

$$\|y_n + y_m - 2x\|^2 = 4 \left\| \frac{y_n + y_m}{2} - x \right\|^2 \geq 4h^2,$$

так как $\{(y_n + y_m)/2\} \in M$. Фиксируем $\varepsilon > 0$ и выберем N такое, что для всех

$$n, m > N \quad \|y_n - x\|^2 < \frac{\varepsilon}{4 + h^2}. \quad \text{Тогда из (1.26) получаем}$$

$$\|y_n - x\|^2 \leq 2 \frac{\varepsilon}{4 + h^2} + 2 \frac{\varepsilon}{4 + h^2} - 4h^2 = \varepsilon.$$

Таким образом, y_n является последовательностью Коши и существует $y \in \mathcal{H}$ такой, что $\|y_n - y\| \rightarrow 0$. Так как M замкнуто, то $y \in M$ и $\|y - x\| = h$ в силу непрерывности скалярного произведения.

Остается показать, что $x - y \perp M$. Предположим, что это так. Тогда существует элемент $w \in M$ такой, что $(x - y, w) = r > 0$. Поэтому для любого c $y + cw \in M$ и в силу (1.25) $\|x - y - cw\| \geq h = \|x - y\|$ и $0 \leq \|x - y - cw\|^2 - \|x - y\|^2 - 2c(x - y, w) + c^2\|w\|^2 - \|x - y\|^2$. Это означает, что $2(x - y, w) = 2r < c\|w\|^2$. Но так как c было произвольным, то наблюдается противоречие при, например, $c = r/\|w\|^2$.

Таким образом, элемент $z = x - y$ должен лежать в M^\perp .

Это завершает доказательство.

Предположим, что имеется случайный процесс $\{X_t, t \in R\}$, определенный на некотором вероятностном пространстве $(\Omega, \mathfrak{F}, P)$ такой, что для каждого t

$$M[X_t] = 0, \quad M[X_t^2] < \infty. \quad (1.27)$$

Процесс, удовлетворяющий условию (1.27), иногда называют процессом *второго порядка*. Каждая случайная величина X_t является элементом пространства H_0 , так что процесс $\{X_t\}$ можно рассматривать, как кривую (однопараметрическое семейство точек) в гильбертовом пространстве. Эта кривая непрерывна, если $\{X_t\}$ непрерывен в среднеквадратическом, т.е. если $M[(X_t - X_s)^2] \rightarrow 0$ при $s \rightarrow t$.

Введем семейство гильбертовых подпространств $\mathcal{H}_t^X \subset \mathcal{H}_0$, соответствующие $\{X_t\}$, где $\mathcal{H}_t^X = \mathcal{L}\{X_s, 0 \leq s \leq t\}$. Оно состоит из всех линейных комбинаций $\sum a_i X_{t_i}$ с $t_i \leq t$ и их пределов в среднем квадратическом. Заметим, что это семейство является невозрастающим $\mathcal{H}_t^X \subset \mathcal{H}_{t_1}^X \subset \mathcal{H}^X = \mathcal{L}\{X_s, s \geq 0\}$ при $t < t_1$.

Так как \mathcal{H}_t^X является подпространством, то любой элемент $Y \in \mathcal{H}_0$ имеет единственное разложение $Y = Y_1 + Y_2$, где $Y_1 \in \mathcal{H}_t^X$ и $Y_2 \perp \mathcal{H}_t^X$. Обозначим \mathcal{P}_t^X (или просто \mathcal{P}_t) оператор проектирования, который переводит каждый элемент в его проекцию на \mathcal{H}_t^X , т.е. $\mathcal{P}_t^X Y = Y_1$. Оператор проектирования обладает следующими свойствами: $(\mathcal{P}_t^X)^2 Y = \mathcal{P}_t^X (\mathcal{P}_t^X) Y = \mathcal{P}_t^X Y$ и при $s < t$

$$\mathcal{P}_s^X Y = \mathcal{P}_s^X \mathcal{P}_t^X Y. \quad (1.28)$$

Действительно, $Y = \mathcal{P}_t Y + (Y - \mathcal{P}_t Y)$ и $(Y - \mathcal{P}_t Y) \perp \mathcal{H}_t^X$. Так как $\mathcal{H}_s^X \subset \mathcal{H}_t^X$, то $(Y - \mathcal{P}_t Y) \perp \mathcal{H}_s^X$. Свойство (1.28) показывает, что можно проектировать по этапам.

Из теоремы 1.3.1 следует, что $\mathcal{P}_t^X Y$ обладает следующим свойством:

$$\|Y - \mathcal{P}_t Y\| = \min_{Z \in \mathcal{H}_t^X} \|Y - Z\|.$$

Так как $\|Y - Z\|^2 = M[(Y - Z)^2]$, то это означает, что $\mathcal{P}_t^X Y$ является наилучшей (в среднем квадратическом смысле) линейной оценкой для Y при заданном $\{X_s, 0 \leq s \leq t\}$. Таким образом, в принципе решена задача линейного оценивания для случайных процессов с непрерывным временем: как и в одномерном случае, наилучшей оценкой случайной величины Y является ее проекция на подпространство, порожденное наблюдениями. Заметим, что для $t > s$ $(Y - \mathcal{P}_t Y) \perp (\mathcal{P}_t Y - \mathcal{P}_s Y)$, так как $(\mathcal{P}_t Y - \mathcal{P}_s Y) \perp \mathcal{H}_s^X$. Поэтому $Y - \mathcal{P}_s Y = (Y - \mathcal{P}_t Y) + (\mathcal{P}_t Y - \mathcal{P}_s Y)$, откуда $\|Y - \mathcal{P}_s Y\|^2 = \|Y - \mathcal{P}_t Y\|^2 + \|\mathcal{P}_t Y - \mathcal{P}_s Y\|^2 \geq \|Y - \mathcal{P}_t Y\|^2$, что выражает тот очевидный факт, что ошибка оценивания $\|Y - \mathcal{P}_t Y\|^2$ уменьшается с ростом t . Если для некоторого t_1 $Y \in \mathcal{H}_{t_1}^X$, то, очевидно, $Y = \mathcal{P}_{t_1} Y$. Интерес представляет случай $Y = X_{t_1}$. Тогда для $t < t_1$ $\mathcal{P}_t Y$ является наилучшим линейным прогнозом для X_{t_1} при заданном $\{X_s, 0 \leq s \leq t\}$. Сходится ли в среднеквадратическом $\mathcal{P}_{t_1} Y$ к X_{t_1} ? Чтобы ответить на этот вопрос необходимо рассмотреть свойства непрерывности процесса $\{X_t\}$.

1.3.2. Непрерывность в среднем квадратическом

1.3.4. Предложение. а) Процесс $\{X_t\}$ непрерывен в среднем квадратическом в точке t , если и только если его ковариационная функция $K(\cdot, \cdot)$ непрерывна в диагональной точке (t, t) .

б) Если процесс $\{X_t\}$ непрерывен в среднем квадратическом при всех t , то $K(\cdot, \cdot)$ непрерывна в каждой точке $(s, t) \in (R)^2$.

Доказательство. Если $K(\cdot, \cdot)$ непрерывна в точке (t, t) то

$$\begin{aligned} M[(X_{t+h} - X_t)^2] &= K(t+h, t+h) - 2K(t+h, t) + K(t, t) = \\ &= \{K(t+h, t+h) - K(t, t)\} - 2\{K(t+h, t) - K(t, t)\} \rightarrow 0 \text{ при } h \rightarrow 0. \end{aligned}$$

Обратно, для любых t, t' можно написать

$$K(t+h, t'+h') - K(t, t') = \{K(t+h, t'+h') - K(t+h, t')\} + \{K(t', t'+h') - K(t', t')\},$$

откуда с помощью неравенства Шварца получаем, что если процесс $\{X_t\}$ непрерывен в среднем квадратическом в точках t и t' , то правая часть равенства стремится к 0 при $h, h' \rightarrow 0$.

Пусть функция $K(t,t)$ непрерывна при всех $t \in R^+$. Тогда процесс $\{X_t\}$ непрерывен в среднем квадратическом при всех $t \in R^+$ и, следовательно, функция $K(s,t)$ непрерывна при всех (s,t) , а не только на диагонали.

Теперь можно ответить на вопрос, поставленный в конце предыдущего раздела. Пусть $\{X_t\}$ - непрерывный в среднем квадратическом процесс и $Y \in \mathcal{H}_t^X$ для некоторого t . Для произвольно заданного $\varepsilon > 0$ можно выбрать

$(s_1, \dots, s_n, a_1, \dots, a_n)$ такие, что $\left\| Y - \sum_{i=1}^n a_i X_{s_i} \right\| < \varepsilon/2$. Для $s_n = t$ можно выбрать

$s_{n-1} \leq t_1 < t$ так, что $\|X_{t_1} - X_t\| < \varepsilon/2 |a_n|$. Но тогда

$$\left\| Y - \left\{ \sum_{i=1}^{n-1} a_i X_{s_i} + a_n X_{t_1} \right\} \right\| \leq \left\| Y - \sum_{i=1}^n a_i X_{s_i} \right\| + \|a_n (X_t - X_{t_1})\|.$$

Обозначим Z выражение слева в скобках. Тогда $Z \in \mathcal{H}_{t_1}^X$, $t_1 < t$ и $\|Y - Z\| < \varepsilon$.

Теперь, так как $\|Y - \mathcal{P}_{t_1}^X Y\| < \|Y - Z\|$, то $\mathcal{P}_{t_1}^X Y \rightarrow Y$ при $t_1 \uparrow t$.

Обычно процесс не бывает непрерывным в среднем квадратическом, если он имеет разрывы в фиксированные моменты времени. Например, рассмотрим процесс $\{X_t\}$ вида

$$X_t(\omega) = \begin{cases} 0, & t < 1, \\ Z(\omega), & t \geq 1, \end{cases}$$

где Z - гауссовская случайная величина с параметрами $(0,1)$. Таким образом, можно представить, что $\mathcal{H}_t^X = \{0\}$ для $t < 1$ и $\mathcal{H}_t^X = \mathcal{L}(Z) = \{aZ : a \in R\}$ для $t \geq 1$.

Тогда, очевидно, $\{X_t\}$ не является непрерывным в среднем квадратическом.

С другой стороны, непрерывность в среднем квадратическом не означает непрерывность выборочных функций процесса. Разрывы в его выборочных функциях происходят в случайные моменты времени, причем вероятность разрыва в интервале (s,t) мала при малых $|t-s|$.

Наконец, заметим, что если $\{X_t\}$ непрерывен в среднем квадратическом, то гильбертово пространство \mathcal{H}^X сепарабельно, так как, если Q означает рациональные числа на R , то $\{X_r, r \in Q\}$ - счетное множество и $\mathcal{H}^X \in \mathcal{L}\{X_r, r \in Q\}$, так как для любого $t \in R$ $X_t = \lim_{r \rightarrow t} X_r, r \in Q$.

1.3.3. Интегралы в среднем квадратическом смысле

Пусть $\{X_t\}$ - непрерывный в среднем квадратическом процесс.

Рассмотрим интеграл вида

$$Y(\omega) = \int_a^b g(t)X_t(\omega) dt. \quad (1.29)$$

Это интеграл Лебега по t от функции $g(\cdot)X(\cdot, \omega)$ при каждом фиксированном ω , порождающим новую случайную величину Y . Так как процесс $\{X_t\}$ был определен как набор случайных величин, определенных при каждом t , то это не гарантирует, что при фиксированном ω и меняющемся t будет получена измеримая функция по t , что требуется в определении интеграла (1.29). Можно заменить случайную величину X_t другой случайной величиной \tilde{X}_t , такой, что $P[X_t = \tilde{X}_t] = 1$. Таким образом, проблема состоит в нахождении такой измеримой по t функции $\tilde{X}(t, \omega)$, обладающей тем свойством, что $P[X(t, \omega) = \tilde{X}(t, \omega) \text{ при всех } t] = 1$. Такой процесс $\{\tilde{X}_t\}$ называется *измеримой версией* процесса $\{X_t\}$.

1.3.5. Предложение. У каждого непрерывного в среднем квадратическом процесса существует измеримая версия.

В соответствии с этим в дальнейшем будут рассматриваться только измеримые процессы и интегралы типа (1.29) будут иметь смысл.

1.3.6. Предложение. Пусть $g: [a, b] \rightarrow R$ - измеримая функция такая, что

$$\int_a^b g^2(s) ds < \infty, \quad (1.30)$$

а $\{X_t\}$ - непрерывный в среднем квадратическом процесс. Тогда случайная величина Y , определенная формулой (1.29), принадлежит к гильбертовому пространству \mathcal{H}^X и

$$\|Y(\omega)\|^2 = \int_a^b \int_a^b g(t) g(s) K(t, s) dt ds, \quad (1.31)$$

где $K(t, s) = \text{cov}[X_t, X_s]$.

Предложение 1.3.6 остается справедливым, если (1.30) заменить более слабым условием

$$\int_a^b |g(s)| ds < \infty.$$

1.3.4. Гильбертово пространство и гауссовские процессы

Ковариационная структура гильбертова подпространства $\mathcal{H}^X \subset \mathcal{H}$, соответствующая процессу второго порядка $\{X_t\}$, полностью определяется ковариационной функцией. Так как в гауссовском случае ковариационная функция характеризует также и распределение процесса, то ясно, что между этими понятиями должна быть связь.

1.3.7. Предложение. Пусть $\{X_n\}$ - последовательность гауссовских случайных величин, сходящаяся в среднем квадратическом к случайной величине X . Тогда X - гауссовская случайная величина.

Это предложение имеет смысл в рамках гильбертова пространства. Действительно, если $\{X_\alpha, \alpha \in I\}$ - произвольное гауссовское семейство элементов $L_2(\Omega, \mathfrak{F}, P)$ (т.е. $(X_{\alpha_1}, \dots, X_{\alpha_N})$ - гауссовский вектор при любом наборе $\alpha_1, \dots, \alpha_N \in I$), то любой элемент из $\mathcal{L}\{X_\alpha, \alpha \in I\}$ - является также гауссовским, поскольку все эти элементы состоят из конечных линейных комбинаций и их пределов в среднем квадратическом.

Пусть теперь $\{X_t\}$ - гауссовский процесс, а Y - гауссовская случайная величина. Пусть также $\hat{Y} = \Gamma_t^X Y$, т.е. \hat{Y} есть проекция Y на \mathcal{H}_t^X . Тогда в силу сделанных замечаний \hat{Y} и $\tilde{Y} = Y - \hat{Y}$ являются также гауссовскими. Так как

$\hat{Y} \perp \tilde{Y}$ (т.е. \hat{Y} и \tilde{Y} некоррелированы), то они независимы. Рассмотрим условную характеристическую функцию $\Phi_{Y|X}(u)$. Она равна

$$\begin{aligned}\Phi_{Y|X}(u) &= M[\{\exp(iuY)\} | X_s, s \leq t] = \\ &= M\left[\left\{\exp(iu\hat{Y})\right\}\left\{\exp(iu\tilde{Y})\right\} | X_s, s \leq t\right] = \left\{\exp(iu\hat{Y})\right\} M[\exp(iu\tilde{Y})].\end{aligned}$$

При переходе к последнему выражению использовалась независимость и то, что $\hat{Y} \in \mathcal{H}_t^X$. Так как \tilde{Y} - гауссовская случайная величина, то $M[\exp(iu\tilde{Y})] = \exp(-\sigma^2 u^2 / 2)$, где $\sigma^2 = D[\tilde{Y}]$. Поэтому условная характеристическая функция будет иметь вид

$$\Phi_{Y|X}(u) = \exp\left(iu\hat{Y} - \sigma^2 u^2 / 2\right).$$

Таким образом, \hat{Y} есть условное среднее Y при заданном $\{X_s, s \leq t\}$, так что \hat{Y} является оценкой для Y при заданном $\{X_s, s \leq t\}$ с минимальной среднеквадратической ошибкой, причем в классе линейных и нелинейных оценок.

При получении этого результата использовались следующие свойства:

1. Для любой случайной величины Z и функции $g(X)$

$$M[g(X)Z | X] = g(X)M[Z | X].$$

2. Если Z не зависит от X , то $M[Z | X] = M[Z]$.

В (1) и (2) X означает $\{X_s, s \leq t\}$.

Еще раз отметим совпадение линейных и нелинейных оценок для гауссовских процессов.

§ 1.4. Гауссовские марковские случайные процессы

Марковский случайный процесс, у которого для всех $t, \tau \in [t_0, T]$ функции $p[x(t)/x(\tau)]$ и $p[x(\tau)]$ являются плотностями гауссовского распределения, называются *гауссовским марковским случайным процессом*.

Так как линейные преобразования гауссовского вектора сохраняют его гауссово свойство, то гауссовский процесс всегда можно представить как

вектор состояния непрерывной линейной динамической системы, возбуждаемой гауссовским чисто случайным процессом и имеющий гауссовский вектор начального состояния

$$\begin{cases} \frac{d}{dt} x(t) = A(t)x(t) + B(t)w(t), \\ x(t_0) = x_0, \end{cases}, \quad (1.32)$$

где $x \in R^n$, $w \in R^r$,

$$M[x(t_0)] = \bar{x}_0, \quad (1.33)$$

$$M\left[\{x(t_0) - \bar{x}(t_0)\}\{x(t_0) - \bar{x}(t_0)\}^T\right] = X(t_0), \quad (1.34)$$

$w(t)$ - гауссовский чисто случайный процесс с

$$M[w(t)] = \bar{w}(t). \quad (1.35)$$

Под понятием «чисто случайный процесс» понимается предельный случай случайного процесса с малым временем корреляции. В качестве примера выберем корреляционную функцию, которая позволит совершить эту предельную операцию

$$M\left[\{w(t) - \bar{w}(t)\}\{w(\tau) - \bar{w}(\tau)\}^T\right] = \chi(t) \exp\left(-\frac{|t - \tau|}{T}\right), \quad (1.36)$$

где $T > 0$ – постоянное число, значительно меньшее, чем постоянные времени линейной динамической системы (1.32), а $\chi(t)$ – матрица ковариаций $w(t)$, т.е.

$$\chi(t) = M\left[\{w(t) - \bar{w}(t)\}\{w(t) - \bar{w}(t)\}^T\right]. \quad (1.37)$$

Таким образом, рассматриваемый гауссовский случайный процесс есть предел гауссовского марковского процесса с «очень большим» вторым моментом и «очень малым» временем корреляции.

Дальше под понятием «чисто случайный процесс» будет использоваться понятие «белый шум», который имеет следующую корреляционную характеристику:

$$M\left[\{w(t) - \bar{w}(t)\}\{w(s) - \bar{w}(s)\}^T\right] = W(t) \delta(t - s). \quad (1.38)$$

Здесь $\delta(t - s) = \delta(\gamma)$ – дельта функция (функция Дирака):

$$\delta(\gamma) = \begin{cases} 0, & \gamma \neq 0, \\ \infty, & \gamma = 0, \end{cases} \quad \int_{-\infty}^{\infty} \delta(\gamma) d\gamma = 1. \quad (1.39)$$

Определим матрицу ковариаций процесса $x(t)$ как

$$X(t) = M \left[\{x(t) - \bar{x}(t)\} \{x(t) - \bar{x}(t)\}^T \right]. \quad (1.40)$$

Среднее значение $x(t)$ определяется математическим ожиданием соотношения (1.32)

$$\frac{d}{dt} \bar{x}(t) = A(t)\bar{x}(t) + B(t)\bar{w}(t), \quad \bar{x}(t_0) - \text{задано}. \quad (1.41)$$

Вычитая (1.41) из (1.32) и умножая затем результат на $\{x(t) - \bar{x}(t)\}^T$, получим

$$\begin{aligned} & \left\{ \frac{d}{dt} \{x(t) - \bar{x}(t)\} \right\} \{x(t) - \bar{x}(t)\}^T = \\ & = A(t) \{x(t) - \bar{x}(t)\} \{x(t) - \bar{x}(t)\}^T + B(t) \{w(t) - \bar{w}(t)\} \{x(t) - \bar{x}(t)\}^T. \end{aligned} \quad (1.42)$$

Если к уравнению (1.42) добавить результат его транспонирования, то придем к

$$\begin{aligned} & \frac{d}{dt} \left\{ \{x(t) - \bar{x}(t)\} \{x(t) - \bar{x}(t)\}^T \right\} = \\ & = A(t) \{x(t) - \bar{x}(t)\} \{x(t) - \bar{x}(t)\}^T + \{x(t) - \bar{x}(t)\} \{x(t) - \bar{x}(t)\}^T A^T(t) + \\ & + B(t) \{w(t) - \bar{w}(t)\} \{x(t) - \bar{x}(t)\}^T + \{x(t) - \bar{x}(t)\} \{w(t) - \bar{w}(t)\}^T B^T(t). \end{aligned} \quad (1.43)$$

Взяв математическое ожидание уравнения (1.43) и используя определение (1.40), найдем, что

$$\begin{aligned} & \frac{d}{dt} X(t) = A(t)X(t) + X(t)A^T(t) + \\ & + B(t)M \left[\{x(t) - \bar{x}(t)\} \{w(t) - \bar{w}(t)\}^T \right] + M \left[\{x(t) - \bar{x}(t)\} \{w(t) - \bar{w}(t)\}^T \right] B^T(t). \end{aligned} \quad (1.44)$$

Пусть $\Omega(t, \tau)$ - фундаментальная матрица решений уравнения (1.32).

Тогда

$$x(t) - \bar{x}(t) = \Omega(t, t_0) [x_0(t) - \bar{x}_0(t)] + \int_{t_0}^t \Omega(t, \tau) B(\tau) [w(\tau) - \bar{w}(\tau)] d\tau. \quad (1.45)$$

Полагаем, что

$$M \left[\{x(t_0) - \bar{x}(t_0)\} \{w(t) - \bar{w}(t)\}^T \right] = 0, \quad (1.46)$$

т.е. случайные отклонения начальных условий не коррелированы со случайным процессом $w(t) - \bar{w}(t)$.

Умножив (1.46) справа на $\{w(t) - \bar{w}(t)\}^T$ и взяв математическое ожидание, получим

$$M \left[\{x(t) - \bar{x}(t)\} \{w(t) - \bar{w}(t)\}^T \right] = \int_{t_0}^t \Omega(t, \tau) B(\tau) M \left[\{w(\tau) - \bar{w}(\tau)\} \{w(\tau) - \bar{w}(\tau)\}^T \right] d\tau. \quad (1.47)$$

Из соотношений (1.44) и (1.47) видно, что вся информация о процессе $w(t) - \bar{w}(t)$, нужная для определения матрицы ковариаций $X(t)$, сводится к знанию взвешенного интеграла от *автокорреляционной матрицы*

$$M \left[\{w(\tau) - \bar{w}(\tau)\} \{w(\tau) - \bar{w}(\tau)\}^T \right]. \quad (1.48)$$

Учитывая (1.38) и (1.39), выражение (1.47) запишется в виде

$$M \left[\{x(t) - \bar{x}(t)\} \{w(t) - \bar{w}(t)\}^T \right] = \int_{t_0}^t \Omega(t, \tau) B(\tau) W(\tau) \delta(t - \tau) d\tau = \frac{1}{2} \Omega(t, t) B(t) W(t).$$

Или, учитывая, что $\Omega(t, t) = I_{r \times r}$ — единичная матрица, получаем:

$$M \left[\{x(t) - \bar{x}(t)\} \{w(t) - \bar{w}(t)\}^T \right] = \frac{1}{2} B(t) W(t). \quad (1.49)$$

Таким образом, учитывая полученный результат, уравнение (1.44) перепишем в виде

$$\begin{cases} \frac{d}{dt} X(t) = A(t)X(t) + X(t)A^T(t) + B(t)W(t)B^T(t), \\ X(t_0) = X_0. \end{cases} \quad (1.50)$$

Уравнения (1.37) и (1.50) являются линейными дифференциальными уравнениями для вектора средних значений $\bar{x}(t)$ и матрицы ковариаций $X(t)$. Эти уравнения друг с другом не связаны, поэтому могут быть вычислены отдельно.

Корреляционная матрица случайного процесса $x(t)$ определяется как

$$K(t, \tau) = M \left[\{x(t) - \bar{x}(t)\} \{x(\tau) - \bar{x}(\tau)\}^T \right]. \quad (1.51)$$

Это детерминированная функция двух переменных t и τ . Матрица ковариаций, определяемая соотношением (1.51), является частным случаем (1.36) $X(t) \equiv K(t, t)$.

Используя (1.01), нетрудно получить следующее соотношение

$$K(t + \tau, t) = \Omega(t + \tau, t) X(t) \quad \text{при } \tau \geq 0. \quad (1.52)$$

Если в (1.50) матрицы A, B, Q содержат постоянные элементы и система (1.32) устойчива, то при $t \rightarrow \infty$ $\frac{d}{dt}X(t) \rightarrow 0$ и, таким образом, матрица $X(t) \rightarrow X$, которая будет определяться решением алгебраического уравнения $AX + XA^T + BWB^T = 0$. (1.53)

Такой процесс называют статистически стационарным. Корреляционные матрицы статистически стационарного процесса являются также стационарными и $K(\tau) = \Omega(\tau)X$ при $\tau \geq 0$.

§ 1.5. Стохастические дифференциальные уравнения

1.5.1. Стохастические интегралы и стохастические дифференциальные уравнения

Решение обыкновенного детерминированного дифференциального уравнения также представляет собой марковский процесс. С помощью винеровского процесса можно построить широкий класс марковских процессов с непрерывными траекториями, определяемых как решение *стохастического дифференциального уравнения* [16, 22, 23, 25, 30] вида

$$\begin{aligned} dx(t) &= f(t, x(t))dt + b(t, x(t))d\xi(t), \quad t \geq 0, \\ x(0) &= x_0, \quad x \in R^n, \quad \xi \in R^r. \end{aligned} \quad (1.54)$$

Здесь x_0 - детерминированный или случайный вектор начальных условий; вектор функция $f(t, x(t)) \in R^n$ и матрица $b(t, x(t))$ размера $n \times r$ заданы; $\xi(t)$ стандартный винеровский процесс.

Уравнение (1.54) является символической записью следующего интегрального тождества:

$$x(t) = x_0 + \int_0^t f(s, x(s))ds + \int_0^t b(s, x(s))d\xi(s). \quad (1.55)$$

Последний интеграл в правой части равенства (1.55) называют *стохастическим*. Возможны различные схемы построения стохастического интеграла. Возьмем произвольное разбиение отрезка $[0, t]$ точками t_i , причем $0 = t_0 < t_1 < t_2 < \dots < t_N = t$. Рассмотрим интегральную сумму

$$\sum_{i=0}^{N-1} b(t_i, x(t_i)) [\xi(t_{i+1}) - \xi(t_i)] = J_H. \quad (1.56)$$

Если при стремлении к нулю длины максимального интервала $t_{i+1} - t_i$ разбиений сумма (1.56) сходится в среднеквадратическом к некоторому пределу, то этот предел называется *стохастическим интегралом Ито* и обозначается

$$\int_0^t b(s, x(s)) d\xi(s). \quad (1.57)$$

При этом сходимость в среднем квадратическом означает, что

$$M \left[\int_0^t b(s, x(s)) d\xi(s) - J_H \right]^2 \rightarrow 0 \text{ при } \max(t_{i+1} - t_i) \rightarrow 0.$$

Для построения интеграла Стратоновича рассмотрим интегральную сумму

$$\sum_{i=0}^{N-1} b \left(t_i, \frac{1}{2} (x(t_{i+1}) + x(t_i)) \right) [\xi(t_{i+1}) - \xi(t_i)] = J_C. \quad (1.58)$$

Предел в среднем квадратическом этой суммы называется *стохастическим интегралом Стратоновича* [24] и обозначается так:

$$\int_0^t b(s, x(s)) \overline{d\xi(s)}. \quad (1.59)$$

Стохастическое уравнение (1.55) понимается соответственно тому, в каком смысле понимается стохастический интеграл в (1.57) или (1.59). Оказывается, что если процесс $x(t)$ удовлетворяет решению *уравнения Стратоновича*

$$\begin{aligned} dx(t) &= f(t, x(t)) dt + b(t, x(t)) \overline{d\xi(t)}, \quad t \geq 0, \\ x(0) &= x_0, \quad x \in R^n, \quad \xi \in R^r. \end{aligned} \quad (1.60)$$

то он также является решением следующего *уравнения Ито*:

$$dx(t) = f(t, x(t)) dt + b(t, x(t)) d\xi(t) + \frac{1}{2} \frac{\partial b(t, x(t))}{\partial x} b(t, x(t)) dt, \quad t \geq 0, \quad (1.61)$$

$$x(0) = x_0, \quad x \in R^n, \quad \xi \in R^r.$$

Ввиду эквивалентности уравнений (1.60) и (1.61) ограничимся рассмотрением стохастических уравнений Ито.

Приведем ряд свойств стохастических интегралов Ито. Определим на основном вероятностном пространстве $(\Omega, \mathfrak{F}, P)$ семейство σ -алгебр F_t , порожденных винеровским процессом $\xi(t)$, удовлетворяющих условиям:

- 1) для любых неотрицательных t и s , $t < s$, справедливо вложение $F_t \subset F_s$;
- 2) винеровский процесс $\xi(t)$ измерим относительно F_t . Последнее означает, что для любого борелевского множества $A \in R^n$ событие $\{\omega: \xi(t) \in A\}$ принадлежит F_t ;
- 3) при любых неотрицательных значениях t и s процесс $\xi(t+s) - \xi(t)$ не зависит от любого из событий σ -алгебры F_t .

Стохастический интеграл Ито $\zeta(t) = \int_0^t a(s) d\xi(s)$ определен для любых

случайных процессов $a(t)$, удовлетворяющих требованиям:

- 1) процесс $a(t)$ измерим относительно F_t при любом t ;
- 2) с вероятностью 1 конечен интеграл

$$J(a) = \int_0^T M[\|a(t)\|^2 / F_0] dt. \quad (1.62)$$

Рассматриваемый как функция верхнего предела стохастический интеграл Ито определяет некоторый случайный процесс $\zeta(t)$ с нулевым

математическим ожиданием $M[\zeta(t)] = M[\zeta(t)/F_0] = M\left[\int_0^t a(s) d\xi(s) / F_0\right] = 0$

и корреляционной функцией $M[\zeta(t_1)\zeta^T(t_2)] = \int_0^{\min(t_1, t_2)} M[a(s)a^T(s) / F_0] ds$.

Значения стохастического интеграла Ито при различных значениях верхнего предела t могут быть согласованы таким образом, что процесс $\zeta(t)$ будет сепарабельным и непрерывным.

1.5.2. Формула Ито

Рассмотрим некоторый процесс $\gamma(t)$, измеримый относительно для любого t , причем

$$\gamma(t_2) - \gamma(t_1) = \int_{t_1}^{t_2} f(t) dt + \int_{t_1}^{t_2} \sigma(t) d\xi(t), \quad 0 \leq t_1 \leq t_2 \leq T.$$

Здесь вектор $f(t)$ и матрица $\sigma(t)$ со случайными элементами измерима относительно F_t при любом t и $J(t) < \infty$, $J(\sigma) < \infty$, где функционал J определен формулой (1.62). Говорят, что сумма $f(t)dt + \sigma(t)d\xi(t)$ есть *стохастический дифференциал* $d\gamma(t)$ процесса $\gamma(t)$, и пишут

$$d\gamma(t) = f(t)dt + \sigma(t)d\xi(t). \quad (1.63)$$

Пусть для скалярной функции $V(t, x)$ существуют непрерывные производные $V_t = \partial V(t, x) / \partial t$, $V_x = \partial V(t, x) / \partial x$, $V_{xx} = \partial^2 V(t, x) / \partial x_i \partial x_j$, $i, j = 1, \dots, n$.

Тогда если процесс $\gamma(t)$ имеет стохастический дифференциал (1.63), то процесс $\eta(t) = V(t, \gamma(t))$ также имеет стохастический дифференциал $d\eta(t)$. Этот дифференциал выражается с помощью *формулы Ито*:

$$d\eta(t) = \left[V_t(t, \gamma(t)) + V_x(t, \gamma(t)) f(t) + \frac{1}{2} \text{Tr} \sigma(t) V_{xx}(t, \gamma(t)) \right] dt + V_x(t, \gamma(t)) \sigma(t) d\xi(t). \quad (1.64)$$

Здесь Tr - след матрицы. Формула Ито представляет собой аналог формулы дифференцирования сложной функции.

1.5.3. Марковские диффузионные процессы

Решение стохастического дифференциального уравнения (1.54) определяется как случайный процесс $x(t)$ измеримый относительно F_t при каждом t и удовлетворяющий интегральному тождеству (1.55) при любом t с вероятностью 1.

Теорема 1.5.1. Пусть функции $f(t, x)$ и $b(t, x)$ в (1.54) измеримы по совокупности аргументов $t \geq 0$, $x \in R^n$ и удовлетворяют условиям

$$\|f(t, x)\|^2 + \|b(t, x)\|^2 \leq C(1 + \|x\|^2), \quad C \geq 0, \quad (1.65)$$

$$\|f(t, x) - f(t, y)\| + \|b(t, x) - b(t, y)\| \leq C\|x - y\| \quad (1.66)$$

для любых $x, y \in R^n$. Тогда существует, и при том единственное, решение уравнения (1.54) на любом отрезке $[0, T]$. При этом единственность понимается в следующем смысле: если $x_1(t)$ и $x_2(t)$ - два непрерывных решения уравнения (1.54), то при любом $T > 0$ справедливо соотношение

$$P\left(\sup_t \|x_1(t) - x_2(t)\| > 0\right) = 0, \quad 0 \leq t \leq T.$$

Решение (1.54) представляет собой марковский процесс, переходная вероятность которого

$$P(t, x, s, A) = P(X_{t,x}(s) \in A). \quad (1.67)$$

В равенстве (1.67) $X_{t,x}(s)$ есть решение уравнения (1.54) при $s > t$ и начальным условием $X_{t,x}(t) = x$. Если вектор x - случайный, то при выполнении условий (1.65) имеет место неравенство

$$M\|X_{t,x}(t_1)\|^4 \leq C_1(1 + M\|x\|^4). \quad (1.68)$$

Здесь C_1 зависит от t, t_1 и постоянной C в условии (1.65).

Во многих задачах возникает необходимость вычисления средних значений некоторых функционалов от решений уравнения (1.54). В ряде случаев эту задачу можно свести к решению краевой задачи для уравнений в частных производных. Пусть, например, требуется вычислить математическое ожидание

$$v(t, x) = M[F(X_{t,s}(s))], \quad s \geq t. \quad (1.69)$$

Здесь s фиксировано, а $F(X_{t,s}(s))$ - заданная непрерывная ограниченная скалярная функция. Значение функционала (1.69) зависит от начального момента t и начального момента $x(0)$.

Предположим, что функции $f(t, x), b(t, x)$ уравнения (1.54) и $F(X_{t,s}(s))$ определены при $0 \leq t \leq T, x \in R^n$. и имеют в этой области непрерывные ограниченные производные по x до второго порядка. Тогда функция $v(t, x)$ имеет непрерывные производные по x до второго порядка включительно, дифференцируема по t и удовлетворяет уравнению

$$Lv(t, x) = 0, \quad t \leq s, \quad x \in R^n. \quad (1.70)$$

Дифференциальный оператор L , называемый *производящим оператором марковского процесса* (1.54), выражается формулой

$$L = \frac{\partial}{\partial t} + f^T(t, x) \frac{\partial}{\partial x(t)} + \frac{1}{2} b(t, x) b^T(t, x) \frac{\partial^2}{\partial x^2(t)}. \quad (1.71)$$

Из (1.69) вытекает начальное условие для функции $v(t, x)$:

$$\lim_{t \rightarrow s} v(t, x) = F(x), \quad x \in R^n. \quad (1.72)$$

Решив задачу Коши (1.70) – (1.72), получим значение функции (1.69) при производных (t, x) . Уравнение (1.70) называется *обратным уравнением Колмогорова*.

1.5.4. Линейные стохастические уравнения

Линейное стохастическое уравнение имеет вид

$$\begin{aligned} dx(t) &= A(t)x(t)dt + \sigma(t)d\xi(t) + \eta(t)dt, \quad t \geq 0, \\ x(0) &= x_0, \quad x \in R^n, \quad \xi \in R^r, \quad \eta \in R^n. \end{aligned} \quad (1.73)$$

Здесь A и σ – заданные матрицы соответствующей размерности с непрерывными элементами, $\eta(t)$ заданная непрерывная функция. Случайная величина $x(0)$ имеет характеристики

$$\begin{aligned} M[x_0] &= m_0, \quad M[(x_0 - m_0)(x_0 - m_0)^T] = D_0, \\ m(t) &= M[x(t)], \quad D(t) = M\left[\{x(t) - m(t)\}\{x(t) - m(t)\}^T\right]. \end{aligned} \quad (1.74)$$

Используя формулу Ито, можно получить следующие уравнения

$$\frac{d}{dt} m(t) = A(t)m(t) + f(t), \quad m(0) = m_0, \quad (1.75)$$

$$\frac{d}{dt} D(t) = A(t)D(t) + D(t)A^T(t) + \sigma(t)\sigma^T(t), \quad D(0) = D_0. \quad (1.76)$$

Пусть $\Omega(t, \tau)$ – фундаментальная матрица решений линейного уравнения

$$\frac{d}{dt} z(t) = A(t)z(t). \quad \text{Тогда в силу (1.75), (1.76) математическое ожидание процесса}$$

(1.73) и матрица ковариации выражается соотношениями

$$\begin{aligned} m(t) &= \Omega(t, 0)m_0 + \int_{t_0}^t \Omega(t, \tau)f(\tau) d\tau, \\ D(t) &= \Omega(t, 0)D_0\Omega^T(t, 0) + \int_{t_0}^t \Omega(t, \tau)\sigma(\tau)\sigma^T(\tau)\Omega^T(t, \tau) d\tau. \end{aligned} \quad (1.77)$$

Аналогично можно записать выражения для моментов решений линейного уравнения с диффузией, зависящей от фазовых координат

$$\begin{aligned} dx(t) &= A(t)x(t)dt + \sigma(t)x(t)d\xi(t) + \eta(t)dt, \quad t \geq 0, \\ x(0) &= x_0, \quad x \in R^n, \quad \xi \in R^r, \quad \eta \in R^1, \end{aligned}$$

где A, σ - матрицы размера $n \times n$ с непрерывными элементами, $\xi(t)$ - стандартный скалярный винеровский процесс. В этом случае уравнение для $m(t)$ записывается в виде (1.75), а уравнение для $D(t)$ - в виде

$$\frac{d}{dt} D(t) = A(t)D(t) + D(t)A^T(t) + \sigma(t) \{ D(t) + m(t)m^T(t) \} \sigma^T(t), \quad D(0) = D_0.$$