

# Условные матожидания и пространство $\mathbb{L}^2$

## 1 Нормальное распределение

Для случайной величины  $\xi \sim \mathcal{N}(m, \sigma^2)$  функция плотности распределения ( $\sigma > 0$ ):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad \mathbb{E}\xi = \text{med}(\xi) = \text{mode}(\xi) = m, \quad \mathbb{V}\xi = \sigma^2$$

Для  $n$ -мерного случайного вектора  $\xi \sim \mathcal{N}(\mathbf{m}, \Sigma)$  функция плотности распределения ( $\Sigma \succ 0$ ):

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m})^T \Sigma^{-1}(\mathbf{x} - \mathbf{m})\right), \quad \mathbb{E}\xi = \text{mode}(\xi) = \mathbf{m}, \quad \mathbb{V}\xi = \Sigma$$

Линейное преобразование нормального случайного вектора:

$$\eta = A\xi + b, \quad \xi \sim \mathcal{N}(\mathbf{m}, \Sigma) \quad \implies \quad \eta \sim \mathcal{N}(A\mathbf{m} + b, A\Sigma A^T)$$

**Опр.** Случайный вектор  $\xi$  имеет *нормальное распределение* с параметрами  $(\mathbf{m}, \Sigma)$ , если его характеристическая функция имеет вид

$$\varphi_\xi(\mathbf{t}) = \exp\left(it^T \mathbf{m} - \frac{1}{2}\mathbf{t}^T \Sigma \mathbf{t}\right)$$

В случае, когда  $\Sigma \succ 0$ , данное определение эквивалентно классическому (через функцию плотности). Оно обобщает его на случай вырожденных матриц  $\Sigma$ .

Свойства:

1. Если  $X$  — нормальный случайный вектор, то  $Y = AX + b$  — тоже. При этом

$$Y = AX + b, \quad X \sim \mathcal{N}(\mathbf{m}, \Sigma) \quad \implies \quad Y \sim \mathcal{N}(A\mathbf{m} + b, A\Sigma A^T)$$

2. Пусть  $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$ ,  $i = \overline{1, n}$  — независимые нормальные случайные величины. Тогда

$$X_1 + \dots + X_n \sim \mathcal{N}(m_1^2 + \dots + m_n^2, \sigma_1^2 + \dots + \sigma_n^2)$$

3. Пусть  $X, Y$  — подвекторы нормального случайного вектора:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{m}_X \\ \mathbf{m}_Y \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}\right)$$

Тогда маргинальные распределения  $X$  и  $Y$  являются нормальными:

$$X \sim \mathcal{N}(\mathbf{m}_X, \Sigma_{XX}), \quad Y \sim \mathcal{N}(\mathbf{m}_Y, \Sigma_{YY}),$$

и условное распределение вида  $Y|X = x$  тоже является нормальным с параметрами  $(\mathbf{m}, \Sigma)$ :

$$\mathbf{m} = \mathbb{E}[Y | X = x] = \mathbf{m}_Y + \Sigma_{YX} \Sigma_{XX}^{-1}(x - \mathbf{m}_X)$$

$$\Sigma = \mathbb{V}[Y | X = x] = \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}$$

Здесь  $\mathbb{E}[Y | X = x]$  и  $\mathbb{V}[Y | X = x]$  — условные матожидания и дисперсия соответственно.

4. Для компонент нормального случайного вектора  $X, Y$ :

$$X, Y \text{ независимы} \quad \iff \quad X, Y \text{ не коррелируют}$$

Просто для нормальных случайных величин это свойство неверно.

## 2 Условные матожидания

Материалы этого и следующего разделов взяты из конспекта семинара 6 и из пособия В.В.Некруткина “Условные математические ожидания относительно  $\sigma$ -алгебр и отображений”, которое можно найти [здесь](#).

Пусть  $(\Omega, \mathcal{F}, \mathbb{P})$  — вероятностное пространство.

**Опр.** Пусть  $\mathcal{A} \subset \mathcal{F}$  —  $\sigma$ -алгебра. Условным матожиданием с.в.  $\xi$  относительно  $\sigma$ -алгебры  $\mathcal{A}$  называется случайная величина  $\mathbb{E}[\xi|\mathcal{A}]$ , такая что:

- $\mathbb{E}[\xi|\mathcal{A}]$  измерима относительно  $\mathcal{A}$ ;
- $\forall B \in \mathcal{A} \rightarrow \int_B \mathbb{E}[\xi|\mathcal{A}](w) d\mathbb{P}(w) = \int_B \xi(w) d\mathbb{P}(w)$ .

Условное матожидание определено с точностью до значений на множестве  $\mathbb{P}$ -меры нуль.

Сама случайная величина  $\xi$  не всегда удовлетворяет этому определению, так как в общем случае  $\xi$  не измерима относительно  $\sigma$ -алгебры  $\mathcal{A}$ .

**Опр.** Условной вероятностью события  $C$  относительно  $\sigma$ -алгебры  $\mathcal{A}$  называется случайная величина

$$\mathbb{P}\{C|\mathcal{A}\} = \mathbb{E}[\mathbb{I}_C|\mathcal{A}]$$

**Опр.**  $\sigma$ -алгеброй, порожденной случайной величиной  $\eta$  называется

$$\mathcal{F}_\eta = \{\eta^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R})\}$$

Так как  $\eta$  —  $\mathcal{F}$ -измеримое отображение, то  $\mathcal{F}_\eta \subset \mathcal{F}$ .

**Утв. 1**  $\xi$   $\mathcal{F}_\eta$ -измерима  $\iff \exists$  борелевская функция  $m: \mathbb{R} \rightarrow \mathbb{R}$ , такая что  $\xi = m(\eta)$ .

**Опр.** Функцией регрессии  $\xi$  на  $\eta$  (в конспекте семинара — условным матожиданием)  $\mathbb{E}[\xi|\eta = y]$  называется такая борелевская функция  $m(y): \mathbb{R} \rightarrow \mathbb{R}$ , что

$$\forall B \in \mathcal{B}(\mathbb{R}) \rightarrow \int_{\eta^{-1}(B)} \xi(w) d\mathbb{P}(w) = \int_B m(y) d\mathbb{P}_\eta(y)$$

Функция  $m(y)$  — та самая функция из утверждения 1. Связь прояснится в следующем определении.

Функция  $m(y)$  определена на  $\mathbb{R}$  с точностью до значений на множестве  $\mathbb{P}_\eta$ -меры нуль, где  $\mathbb{P}_\eta(B) = \mathbb{P}\{\eta \in B\}$  — распределение случайной величины  $\eta$ .

**Опр.** Условным матожиданием  $\xi$  относительно  $\eta$  называется случайная величина  $\mathbb{E}[\xi|\eta]$ , равная

- (а)  $m(\eta)$ , где  $m(y) = \mathbb{E}[\xi|\eta = y]$ ;
- (б)  $\mathbb{E}[\xi|\mathcal{F}_\eta]$ .

Определения (а) и (б) условного матожидания  $\mathbb{E}[\xi|\eta]$  эквивалентны.

Объясним связь этого определения и утверждения 1.  $\mathbb{E}[\xi|\eta]$  —  $\mathcal{F}_\eta$ -измеримая функция по определению (б), поэтому функцию в определении (а) можно взять из утверждения 1.

Итак, важно понимать, что условное матожидание относительно случайной величины — частный случай условного матожидания относительно  $\sigma$ -алгебры. Выпишем полезные свойства этих условных матожиданий.

Свойства: ( $\xi_1, \xi_2, \eta$  — с.в. с конечными матожиданиями)

1. Условное матожидание единственно с точностью до значений на множестве  $\mathbb{P}$ -меры нуль;
2.  $\mathbb{E}[a\xi_1 + b\xi_2 | \eta] = a\mathbb{E}[\xi_1|\eta] + b\mathbb{E}[\xi_2|\eta]$ ;
3. Если  $\xi_1 \leq \xi_2$  п.в., то  $\mathbb{E}[\xi_1|\eta] \leq \mathbb{E}[\xi_2|\eta]$  п.в.;

4. Неравенство Йенсена: если  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  — выпуклая борелевская функция, а  $\mathbb{E}|\varphi(\xi)| < \infty$ , то

$$\mathbb{E}[\varphi(\xi) \mid \mathcal{A}] \geq \varphi(\mathbb{E}[\xi \mid \mathcal{A}]) \quad \mathbb{P}\text{-почти всюду на } \Omega$$

5. Если  $\xi$  измерима относительно  $\mathcal{A}$  и  $\mathbb{E}|\xi\eta| < \infty$ , то  $\mathbb{E}[\xi\eta \mid \mathcal{A}] = \xi\mathbb{E}[\eta \mid \mathcal{A}]$  п.в.;

6. Формулы полного матожидания и полной вероятности:

$$\boxed{\mathbb{E}\xi = \mathbb{E}[\mathbb{E}[\xi|\eta]], \quad \mathbb{P}\{B\} = \mathbb{E}[\mathbb{P}\{B|\mathcal{A}\}]}$$

7. Если  $\xi$  и  $\eta$  — независимые с.в., то  $\mathbb{E}[\xi|\eta] = \mathbb{E}\xi$  п.в.;

8.  $\mathbb{E}[\xi|\xi] = \xi$  п.в.

Заметим, что пока никаких содержательных утверждений мы не получили. Перечисленные выше свойства вытекают напрямую из определения. Пока вообще непонятно, как на практике искать  $\mathbb{E}[\xi|\eta]$ . Это обсудим чуть ниже.

## Условное матожидание относительно случайного вектора

Пусть  $\vec{\eta} : \Omega \rightarrow \mathbb{R}^n$ ,  $\vec{\eta} = (\eta_1, \dots, \eta_n)^T$  — случайный вектор.

Все определения, утверждения и свойства переносятся.

**Опр.**  $\sigma$ -алгеброй, порожденной случайным вектором  $\vec{\eta}$  называется

$$\mathcal{F}_{\vec{\eta}} = \{\vec{\eta}^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R}^n)\}$$

Как и ранее, так как  $\vec{\eta}$  — отображение, измеримое относительно пары  $(\mathcal{F}, \mathcal{B}(\mathbb{R}^n))$ , то  $\mathcal{F}_{\vec{\eta}} \subset \mathcal{F}$ .

**Опр.** Функцией регрессии  $\xi$  на  $\vec{\eta}$   $\mathbb{E}[\xi|\vec{\eta} = \vec{y}]$  называется такая борелевская функция  $m(\vec{y}) : \mathbb{R}^n \rightarrow \mathbb{R}$ , что

$$\forall B \in \mathcal{B}(\mathbb{R}^n) \rightarrow \int_{\vec{\eta}^{-1}(B)} \xi(w) d\mathbb{P}(w) = \int_B m(\vec{y}) d\mathbb{P}_{\vec{\eta}}(\vec{y})$$

**Опр.** Условным матожиданием  $\xi$  относительно  $\vec{\eta}$  называется случайная величина  $\mathbb{E}[\xi \mid \vec{\eta}]$ , равная

(а)  $m(\vec{\eta})$ , где  $m(\vec{y}) = \mathbb{E}[\xi|\vec{\eta} = \vec{y}]$ ;

(б)  $\mathbb{E}[\xi \mid \mathcal{F}_{\vec{\eta}}]$ .

Отметим, что условное матожидание относительно случайного вектора — это именно случайная величина, а не случайный вектор.

**Утв. 1'**  $\xi \in \mathcal{F}_{\vec{\eta}} \iff \exists$  борелевская функция  $m : \mathbb{R}^n \rightarrow \mathbb{R}$ , такая что  $\xi = m(\vec{\eta})$ .

## Связь с интуитивным определением

Зададимся вопросом, как найти  $\mathbb{E}[\xi \mid A]$  — условное матожидание относительно события  $A \subset \Omega$ , где  $\xi$  — случайная величина. Интуитивно понятно, что это число должно являться в некотором смысле средним значением с.в.  $\xi$  на множестве  $A$ . Поэтому естественно ожидать следующее выражение:

$$\mathbb{E}[\xi \mid A] \approx \sum_i x_i \cdot \frac{\mathbb{P}\{D_i\}}{\mathbb{P}\{A\}}, \quad (*)$$

где  $A = \bigcup_i D_i$ ,  $x_i$  — значения  $\xi$  на  $D_i$ . То есть мы разбиваем  $A$  на подмножества, на которых  $\xi$  постоянна, и умножаем на вероятность попасть в эти подмножества (при условии, что мы находимся в  $A$ ).

Проведем строгие рассуждения. Заметим, что нашу задачу можно переформулировать через известные термины:

$$\mathbb{E}[\xi \mid A] = \mathbb{E}[\xi \mid \mathbb{I}_A = 1] = m(1),$$

где  $m$  — функция регрессии  $\xi$  на  $\mathbb{I}_A$ .

В определение функции  $m(y)$  подставим множество  $B = \{1\}$ . Тогда  $\mathbb{I}_A^{-1}(1) = A$ , поэтому

$$\int_A \xi(w) d\mathbb{P}(w) = \int_{\{1\}} m(y) d\mathbb{P}_{\mathbb{I}_A}(y) = m(1) \cdot \mathbb{P}_{\mathbb{I}_A}\{1\} = m(1) \cdot \mathbb{P}\{A\}$$

Отсюда получаем, что если  $\mathbb{P}\{A\} > 0$ , то

$$\mathbb{E}[\xi \mid A] = \frac{1}{\mathbb{P}\{A\}} \int_A \xi(w) d\mathbb{P}(w)$$

Это выражение соответствует формуле (\*), которую мы и ожидали получить. Это выражение можно записать в следующем виде:

**Опр.** Условным матожиданием с.в.  $\xi$  относительно события  $A$  ( $\mathbb{P}\{A\} > 0$ ) называется

$$\mathbb{E}[\xi \mid A] = \frac{\mathbb{E}[\xi \cdot \mathbb{I}_A]}{\mathbb{P}\{A\}}$$

На самом деле это скорее утверждение, которое мы доказали, нежели определение.

## Формулы полной вероятности и полного матожидания

Пусть  $\xi$  — случайная величина, а  $\mathcal{A} \subset \mathcal{F}$  —  $\sigma$ -алгебра. Тогда положив в определении  $\mathbb{E}[\xi \mid \mathcal{A}]$  множество  $B = \Omega$ , мгновенно получаем формулу полного матожидания

$$\int_{\Omega} \mathbb{E}[\xi \mid \mathcal{A}](w) d\mathbb{P}(w) = \int_{\Omega} \xi(w) d\mathbb{P}(w) \quad \Longleftrightarrow \quad \mathbb{E}[\mathbb{E}[\xi \mid \mathcal{A}]] = \mathbb{E}\xi$$

Аналогично получается формула полной вероятности. Однако возникает вопрос, а что вообще такое  $\mathbb{E}[\xi \mid \mathcal{A}]$ ? Как его искать?

Исследуем структуру  $\mathbb{E}[\xi \mid \mathcal{A}]$  и формулы полной вероятности в некоторых частных случаях.

**Утв. 2** (случай дискретной случайной величины)

Пусть  $\xi$  — произвольная с.в., а  $\eta$  — дискретная случайная величина со значениями  $\{x_k\}$ . Тогда справедлива формула полного матожидания:

$$\mathbb{E}\xi = \sum_k \mathbb{E}[\xi \mid \eta = x_k] \cdot \mathbb{P}\{\eta = x_k\}$$

и для произвольного события  $A \in \mathcal{F}$  справедлива формула полной вероятности:

$$\mathbb{P}\{A\} = \sum_k \mathbb{P}\{A \mid \eta = x_k\} \cdot \mathbb{P}\{\eta = x_k\}$$

Не приводя строгого доказательства, объясним почему так происходит. Сигма-алгеброй в этом случае является

$$\mathcal{F}_{\eta} = \sigma(\eta^{-1}(x_1), \eta^{-1}(x_2), \dots),$$

то есть  $\sigma$ -алгебра порождена не более чем счетным разбиением  $\Omega = \bigcup_k \{\eta = x_k\}$ . Кроме того, по определению  $\mathbb{E}[\xi \mid \mathcal{F}_{\eta}]$  измерима относительно  $\mathcal{F}_{\eta}$ , поэтому (см. упражнение 2 из конспекта семинара) условное матожидание представимо в виде

$$\mathbb{E}[\xi \mid \mathcal{F}_{\eta}] = \sum_k z_k \mathbb{I}_{\{\eta = x_k\}}, \quad z_k = \mathbb{E}[\xi \mid \eta = x_k]$$

Можно доказать, что числа  $z_k$  имеют такой вид. Далее по свойству линейности матожидания получаем конечное выражение.

**Утв. 3** (случай непрерывной случайной величины)

Пусть  $\xi, \eta$  — непрерывные с.в. с плотностями  $f_\xi, f_\eta$ . Тогда справедлива *формула полного матожидания*:

$$\mathbb{E}\xi = \int_{-\infty}^{+\infty} \mathbb{E}[\xi \mid \eta = x] f_\eta(x) dx$$

и для события  $A \in \mathcal{F}$  справедлива *формула полной вероятности*:

$$\mathbb{P}\{A\} = \int_{-\infty}^{+\infty} \mathbb{P}\{A \mid \eta = x\} f_\eta(x) dx$$

**Утв. 4** (теорема Байеса)

Для непрерывных случайных величин  $\xi$  и  $\eta$  справедлива *формула Байеса*:

$$f_\xi(x) = \int_{-\infty}^{+\infty} f_{\xi|\eta}(x \mid y) f_\eta(y) dy$$

Заметим, что последнее равенство легко получить напрямую. Подынтегральное выражение равно совместной плотности распределения  $f_{(\xi, \eta)}(x, y)$ , а его интегрирование как раз дает маргинальное распределение.

Эти формулы удобно применять при решении задач, потому что часто условные матожидания вида  $\mathbb{E}[\xi \mid \eta = y]$  проще находить, чем просто  $\mathbb{E}\xi$  (например, когда  $\xi$  явно зависит от  $\eta$ , можно заменить  $\eta$  на  $y$  и убрать условие).

### 3 Пространство $\mathbb{L}^2$

**Опр.** Пусть  $(\Omega, \mathcal{F}, \mathbb{P})$  — вероятностное пространство. Обозначим через  $L^2 \equiv L^2(\Omega, \mathcal{F}, \mathbb{P})$  множество всех случайных величин на  $\Omega$  с конечным вторым моментом:

$$L^2 = \{\text{с.в. } \xi : \Omega \rightarrow \mathbb{R} \mid \mathbb{E}[\xi^2] < +\infty\}$$

**Утв. 5**  $L^2$  — линейное пространство (относительно сложения с.в. и умножения на вещественное число).

Введем на  $L^2$  функцию:

$$\langle \cdot, \cdot \rangle : L^2 \times L^2 \rightarrow \mathbb{R}, \quad \langle \xi, \eta \rangle = \mathbb{E}[\xi\eta]$$

Наша цель: ввести на  $L^2$  скалярное произведение, то есть функцию, удовлетворяющую аксиомам:

1.  $\langle \xi, \eta \rangle = \langle \eta, \xi \rangle \quad \forall \xi, \eta \in L^2$
2.  $\langle \xi_1 + \xi_2, \eta \rangle = \langle \xi_1, \eta \rangle + \langle \xi_2, \eta \rangle \quad \forall \xi_1, \xi_2, \eta \in L^2$   
 $\langle \lambda \xi, \eta \rangle = \lambda \langle \xi, \eta \rangle \quad \forall \lambda \in \mathbb{R} \quad \forall \xi, \eta \in L^2$
3.  $\langle \xi, \xi \rangle > 0 \quad \forall \xi \neq 0$

**Утв. 6** Функция  $\langle \xi, \eta \rangle = \mathbb{E}[\xi\eta]$  удовлетворяет аксиомам 1 и 2 скалярного произведения.

С аксиомой 3 возникает проблема:

$$\langle \xi, \xi \rangle = 0 \iff \mathbb{E}[\xi^2] = 0 \iff \mathbb{P}\{\xi = 0\} = 1 \iff \xi(w) = 0 \text{ п.в. на } \Omega,$$

то есть существуют случайные величины  $\xi \neq 0$ , для которых  $\langle \xi, \xi \rangle = 0$ .

Чтобы обойти эту проблему, будем рассматривать классы эквивалентности случайных величин относительно равенства почти всюду:

$$[\xi] = \{\eta \in L^2 \mid \xi = \eta \text{ п.в. на } \Omega\}$$

Отметим, что равенство почти всюду означает, что  $\mathbb{P}\{\xi \neq \eta\} = 0$ , то есть мера множества на, котором  $\xi$  и  $\eta$  различны, равна 0.

**Опр.** Множество классов эквивалентности случайных величин ( $\mathbb{E}[\xi^2] < \infty$ ) относительно равенства почти всюду называется *пространством  $\mathbb{L}^2$* .

Далее будем отождествлять случайные величины  $\xi$  и их классы эквивалентности  $[\xi]$ .

Заметим, что функция  $\langle \xi, \eta \rangle = \mathbb{E}[\xi\eta]$  не зависит от значений  $\xi$  и  $\eta$  на множестве меры 0 (так как матожидание — интеграл Лебега по мере  $\mathbb{P}$ ). Значит, определим

$$\langle \cdot, \cdot \rangle : \mathbb{L}^2 \times \mathbb{L}^2 \rightarrow \mathbb{R}, \quad \langle [\xi], [\eta] \rangle = \langle \xi, \eta \rangle = \mathbb{E}[\xi\eta],$$

то есть скалярным произведением двух классов эквивалентности назовем матожидание произведения любых двух с.в. из этих классов.

Таким образом, пространство  $\mathbb{L}^2$  с функцией  $\langle \cdot, \cdot \rangle$  является *предгильбертовым* пространством.

Предгильбертово пространство — линейное пространство со скалярным произведением (inner product space).

**Утв. 7**  $\mathbb{L}^2$  — гильбертово пространство.

Гильбертово пространство — полное предгильбертово пространство. Другими словами, скалярное произведение порождает норму, которая порождает метрику:

$$\|\xi\| = \sqrt{\langle \xi, \xi \rangle}, \quad \rho(\xi, \eta) = \|\xi - \eta\|,$$

то есть предгильбертово пространство является метрическим. Под полнотой понимается полнота метрического пространства с порожденной метрикой.

**Утв. 8** Пусть  $\{\xi_n\}, \{\eta_n\} \subset L^2$  — последовательности случайных величин таких, что

$$\mathbb{E}[(\xi_n - \xi)^2] \rightarrow 0, \quad \mathbb{E}[(\eta_n - \eta)^2] \rightarrow 0$$

Тогда

1.  $\xi, \eta \in L^2$ ;
2.  $\mathbb{E}[\xi_n] \rightarrow \mathbb{E}[\xi]$ ;
3.  $\mathbb{E}[\xi_n \eta_n] \rightarrow \mathbb{E}[\xi \eta]$ ;
4.  $\text{cov}(\xi_n, \eta_n) \rightarrow \text{cov}(\xi, \eta)$ ;
5.  $\mathbb{V}[\xi_n] \rightarrow \mathbb{V}[\xi]$ .

**Неравенство Коши-Буняковского.** Для любых  $\xi, \eta \in L^2$ :

$$\mathbb{E}[\xi \eta] \leq \sqrt{\mathbb{E}\xi^2 \cdot \mathbb{E}\eta^2}$$

в частности, для центрированных случайных величин:

$$\text{cov}(\xi, \eta) \leq \sqrt{\mathbb{V}\xi \cdot \mathbb{V}\eta}$$

**Опр.** С.в.  $\xi, \eta \in L^2$  называются *ортгоналными* ( $\xi \perp \eta$ ), если  $\langle \xi, \eta \rangle = \mathbb{E}[\xi \eta] = 0$ .

**Опр.** Проекцией  $\xi$  на  $\eta$  называется  $\pi_\eta(\xi) = \frac{\langle \xi, \eta \rangle}{\|\eta\|^2} \cdot \eta$ .

## Условное матожидание в $\mathbb{L}^2$

Пусть  $H$  — произвольное гильбертово пространство,  $M \subset H$  — его подпространство. Тогда

$$N = M^\perp = \{y \mid y \perp x \ \forall x \in M\} \text{ — ортогональное дополнение}$$

При этом  $H$  раскладывается в *прямую сумму*  $H = M \oplus N$ , и  $\forall z \in H \ \exists$  единственное разложение  $z = x + y$ , где  $x \in M$ ,  $y \in N$  ( $x, y$  — проекции  $z$  на  $M, N$  соответственно).

Пусть  $\mathcal{A} \subset \mathcal{F}$  —  $\sigma$ -алгебра. Изучим, какой вид имеет условное матожидание  $\mathbb{E}[\xi | \mathcal{A}]$ , когда  $\xi \in L^2$ .

**Утв. 9** Если  $\xi \in L^2$ , то  $\mathbb{E}[\xi | \mathcal{A}] \in \mathbb{L}^2$  и единственно.

*Доказательство.*

Функция  $\varphi(x) = x^2$  выпукла, тогда по неравенству Йенсена:

$$(\mathbb{E}[\xi | \mathcal{A}])^2 \leq \mathbb{E}[\xi^2 | \mathcal{A}] \quad \text{п.в.}$$

Беря матожидание от обеих частей неравенства, получаем по формуле полного матожидания:

$$\mathbb{E}[(\mathbb{E}[\xi | \mathcal{A}])^2] \leq \mathbb{E}\xi^2 < \infty \quad \implies \quad \mathbb{E}[\xi | \mathcal{A}] \in L^2$$

Единственность в  $\mathbb{L}^2$  получаем из того, что условное матожидание единственно с точностью до значений на множестве меры нуль.  $\square$

Рассмотрим множество  $M_{\mathcal{A}} \subset \mathcal{F}$ , состоящее из всех с.в., измеримых относительно  $\mathcal{A}$ . Сумма измеримых функций — измеримая функция, измеримая функция, умноженная на число — тоже, поэтому  $M_{\mathcal{A}}$  — замкнутое подпространство гильбертового пространства  $\mathbb{L}^2$ .

Значит, можно находить проекцию на  $M_{\mathcal{A}}$  и строить к нему ортогональное дополнение.

**Утв. 10** Если  $\xi \in L^2$ , то  $\mathbb{E}[\xi | \mathcal{A}]$  — проекция  $\xi$  на подпространство  $M_{\mathcal{A}}$ .

*Доказательство.*

По общему свойству проекции,  $\xi' = \mathbb{E}[\xi | \mathcal{A}]$  — проекция  $\xi$  на  $M_{\mathcal{A}}$   $\iff \xi - \xi' \in M_{\mathcal{A}}^\perp \iff$

$$\forall \eta \in M_{\mathcal{A}} \rightarrow \langle \xi - \xi', \eta \rangle = 0$$

Проверим последнее равенство. По линейности и свойству 5 условного матожидания ( $\eta$  измерима отн.  $\mathcal{A}$ , поэтому ее можно занести под знак матожидания во втором равенстве):

$$\mathbb{E}[(\xi - \mathbb{E}[\xi | \mathcal{A}])\eta] = \mathbb{E}[\xi\eta] - \mathbb{E}[\mathbb{E}[\xi | \mathcal{A}]\eta] = \mathbb{E}[\xi\eta] - \mathbb{E}[\mathbb{E}[\xi\eta | \mathcal{A}]] = \mathbb{E}[\xi\eta] - \mathbb{E}[\xi\eta] = 0$$

Предпоследнее равенство получено по формуле полного матожидания.  $\square$

Как интерпретировать полученный результат? Нетрудно доказать, что

$$\mathbb{E}[(\xi - \mathbb{E}[\xi | \mathcal{A}])^2] = \min_{\eta \in M_{\mathcal{A}}} \mathbb{E}[(\xi - \eta)^2]$$

Таким образом, условное матожидание  $\mathbb{E}[\xi | \mathcal{A}]$  — наилучшее приближение случайной величины в “менее богатой”  $\sigma$ -алгебре по среднеквадратичной метрике.

## Пространство $\mathbb{L}_0^2$

**Опр.** Обозначим через  $L_0^2$  множество всех центрированных случайных величин с конечным вторым моментом:

$$L_0^2 = \{\xi \in L^2 \mid \mathbb{E}\xi = 0\}$$

Скалярное произведение в  $L_0^2$  является ковариацией:

$$\langle \xi, \eta \rangle = \mathbb{E}[\xi\eta] = \text{cov}(\xi, \eta), \quad \xi, \eta \in L_0^2$$

Что будет, если попробовать задать так скалярное произведение на всем пространстве  $L^2$ ?

**Утв. 11** Функция  $\text{cov}(\xi, \eta)$  удовлетворяет аксиомам 1 и 2 скалярного произведения на  $L^2$ .

Рассмотрим аксиому 3:

$$\text{cov}(\xi, \xi) = \mathbb{V}\xi = 0 \quad \implies \quad \xi \equiv \text{const} \text{ п.в.}$$

Поэтому, если рассмотреть в  $L^2$  следующие классы эквивалентности:

$$[[\xi]] = \{\eta \in L^2 \mid \xi - \eta \equiv \text{const} \text{ п.в.}\}, \quad (*)$$

то на полученном фактор-пространстве ковариация уже будет скалярным произведением.

Обозначение  $[\xi]$  сохраним для классов эквивалентности относительно равенства почти всюду.

**Опр.** Множество классов эквивалентности  $(*)$  в  $L^2$  со скалярным произведением  $\text{cov}(\xi, \eta)$  называется *пространством  $\mathbb{L}_0^2$* .

Отметим, что  $\mathbb{L}_0^2$  — подпространство  $\mathbb{L}^2$  в следующем смысле. Рассмотрим в  $\mathbb{L}^2$  подпространство  $M$ , факторизовав множество  $L_0^2$  по равенству почти всюду. Тогда из  $\mathbb{L}_0^2$  в  $M$  легко построить изоморфизм:

$$\varphi : [[\xi]] \mapsto [\xi - \mathbb{E}\xi]$$

*Изоморфизм* — взаимно однозначное соответствие, сохраняющее скалярное произведение.

Несложно убедиться, что  $M$  — замкнутое подпространство  $\mathbb{L}^2$ , а  $\mathbb{L}^2$  — полно. Значит, и  $M$  полно. Изоморфизм сохраняет полноту, поэтому  $\mathbb{L}_0^2$  полно.

**Утв. 12**  $\mathbb{L}_0^2$  — гильбертово пространство.

Как и в  $\mathbb{L}^2$ , можно для удобства отождествлять случайные величины и их классы эквивалентности в  $\mathbb{L}_0^2$ .



## 4 Примеры

### Пример 1 (формула свертки)

Пусть  $\xi, \eta$  — непрерывные случайные величины с плотностями  $f_\xi, f_\eta$ . Найти плотность  $\xi + \eta$ .

**Решение:**

Найдем функцию распределения. По формуле полной вероятности:

$$F_{\xi+\eta}(x) = \mathbb{P}\{\xi + \eta < x\} = \int_{-\infty}^{+\infty} \mathbb{P}\{\xi + \eta < x \mid \eta = y\} f_\eta(y) dy = \int_{-\infty}^{+\infty} \mathbb{P}\{\xi < x - y\} f_\eta(y) dy = \int_{-\infty}^{+\infty} F_\xi(x - y) f_\eta(y) dy$$

Дифференцируя по  $x$  находим плотность (применяем теорему о дифференцировании несобственного интеграла по параметру):

$$f_{\xi+\eta}(x) = \int_{-\infty}^{+\infty} f_\xi(x - y) f_\eta(y) dy$$

Данная формула называется *формулой свертки*.

### Пример 2 (нормальные случайные величины)

Пусть  $(X, Y)$  — нормальный случайный вектор с распределением:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m_X \\ m_Y \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

Найти  $\mathbb{E}[X \mid Y = \alpha]$ ,  $\alpha \in \mathbb{R}$ .

**Решение:**

Рассмотрим  $X, Y$  как элементы пространства  $\mathbb{L}_0^2$ . Представим  $X = X_{\parallel} + X_{\perp}$ , где

$$X_{\parallel} = \pi_Y(X) = \frac{\text{cov}(X, Y)}{\text{cov}(Y, Y)} \cdot Y = \frac{\rho}{1} Y = \rho Y \quad \text{— проекция } X \text{ на } Y \text{ в } \mathbb{L}_0^2$$

$$X_{\perp} = X - X_{\parallel} = X - \rho Y$$

По построению,  $\text{cov}(X_{\perp}, Y) = 0$ , значит, с.в.  $X_{\perp}$  и  $Y$  не коррелируют.  $X_{\perp}$  — линейная комбинация нормально распределенных с.в., значит,  $X_{\perp}$  — нормальная с.в. Это компоненты некоторого нормального случайного вектора, а для них из некоррелированности следует независимость. Значит,  $X_{\perp}$  и  $Y$  независимы.

В силу линейности условного математического ожидания:

$$\mathbb{E}[X \mid Y = \alpha] = \mathbb{E}[X_{\parallel} \mid Y = \alpha] + \mathbb{E}[X_{\perp} \mid Y = \alpha] = \mathbb{E}[\rho Y \mid Y = \alpha] + \mathbb{E}[X_{\perp}] = \alpha\rho + m_X - \rho m_Y$$

Заметим, что полученный результат согласуется с общей формулой для условного распределения компонент нормального случайного вектора.