

# Теоретические основы информатики (концептуальные модели и математические основы)

## Лекция № 7. Радемахеровская сложность

А.С. Шундеев

# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности

# Содержание

## 1 Радемахеровская сложность

- Радемахеровское среднее
- Радемахеровская сложность

## 2 Эмпирические и предсказывающие оценки

## 3 Оценка сверху для радемахеровской сложности

# Радемахеровская сложность

В данной лекции продолжается изучение равномерной сходимости эмпирического риска, но уже с более общих позиций.

# Радемахеровская сложность

В данной лекции продолжается изучение равномерной сходимости эмпирического риска, но уже с более общих позиций.

В частности, изначально не делается никаких дополнительных предположений относительно множества меток  $Y$  класса гипотез  $\mathcal{H} \subseteq Y^X$  и используемой функции потерь.

# Радемахеровская сложность

В данной лекции продолжается изучение равномерной сходимости эмпирического риска, но уже с более общих позиций.

В частности, изначально не делается никаких дополнительных предположений относительно множества меток  $Y$  класса гипотез  $\mathcal{H} \subseteq Y^X$  и используемой функции потерь.

С этой целью вводится понятие **радемахеровской сложности** класса функции и излагается основанный на этом понятии подход к получению так называемых **эмпирических** и **предсказывающих оценок**.

# Радемахеровская сложность

После этого вновь будет рассматриваться задача бинарной классификации, но уже с использованием понятия и свойств радемахеровской сложности.

# Радемахеровская сложность

После этого вновь будет рассматриваться задача бинарной классификации, но уже с использованием понятия и свойств радемахеровской сложности.

Основным результатом здесь является доказательство теоремы Дадли, одним из следствий которой является улучшение оценки для функции сложности обучающей выборки, позволяющей убрать логарифм в её правой части:

$$n_{\mathcal{H}}^{\text{uc}}(\varepsilon, \delta) = \mathcal{O}\left(\frac{\text{vc}(\mathcal{H}) + \ln(1/\delta)}{\varepsilon^2}\right). \quad (1)$$



# Содержание

- 1 Радемахеровская сложность
  - Радемахеровское среднее
  - Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности

# Радемахеровская сложность

## Радемахеровское среднее

В основе определения радемахеровской сложности класса функций лежит понятие радемахеровского среднего ограниченного множества в  $\mathbb{R}^n$ .

# Радемахеровская сложность

## Радемахеровское среднее

В основе определения радемахеровской сложности класса функций лежит понятие радемахеровского среднего ограниченного множества в  $\mathbb{R}^n$ .

С определения этого понятия мы и начнём.

# Радемахеровская сложность

## Радемахеровское среднее

В основе определения радемахеровской сложности класса функций лежит понятие радемахеровского среднего ограниченного множества в  $\mathbb{R}^n$ .

С определения этого понятия мы и начнём.

### Определение 5.7.

**Радемахеровским средним** ограниченного множества  $A \subset \mathbb{R}^n$  ( $n \in \mathbb{N}$ ) называется число

$$\mathcal{R}_n(A) := \mathbf{E} \left[ \sup_{(a_1, \dots, a_n) \in A} \frac{1}{n} \sum_{i=1}^n \varepsilon_i a_i \right],$$

где  $\varepsilon_1, \dots, \varepsilon_n$  – независимые радемахеровские случайные величины.

# Радемахеровская сложность

## Радемахеровское среднее

Следующие свойства непосредственно вытекают из приведённого определения.

# Радемахеровская сложность

## Радемахеровское среднее

Следующие свойства непосредственно вытекают из приведённого определения.

### Утверждение 5.7.

Пусть  $n \in \mathbb{N}$ ,  $A, B \subset \mathbb{R}^n$  – ограниченные множества и  $c \in \mathbb{R}$ .

# Радемахеровская сложность

## Радемахеровское среднее

Следующие свойства непосредственно вытекают из приведённого определения.

### Утверждение 5.7.

Пусть  $n \in \mathbb{N}$ ,  $A, B \subset \mathbb{R}^n$  – ограниченные множества и  $c \in \mathbb{R}$ .

Тогда

$$\textcircled{1} \quad \mathcal{R}_n(cA) = |c| \mathcal{R}_n(A);$$

# Радемахеровская сложность

## Радемахеровское среднее

Следующие свойства непосредственно вытекают из приведённого определения.

### Утверждение 5.7.

Пусть  $n \in \mathbb{N}$ ,  $A, B \subset \mathbb{R}^n$  – ограниченные множества и  $c \in \mathbb{R}$ .

Тогда

- ❶  $\mathcal{R}_n(cA) = |c|\mathcal{R}_n(A)$ ;
- ❷ если  $A \subseteq B$ , то  $\mathcal{R}_n(A) \leq \mathcal{R}_n(B)$ ;



# Радемахеровская сложность

## Радемахеровское среднее

Следующие свойства непосредственно вытекают из приведённого определения.

### Утверждение 5.7.

Пусть  $n \in \mathbb{N}$ ,  $A, B \subset \mathbb{R}^n$  – ограниченные множества и  $c \in \mathbb{R}$ .

Тогда

- 1  $\mathcal{R}_n(cA) = |c|\mathcal{R}_n(A)$ ;
- 2 если  $A \subseteq B$ , то  $\mathcal{R}_n(A) \leq \mathcal{R}_n(B)$ ;
- 3  $\mathcal{R}_n(A + B) = \mathcal{R}_n(A) + \mathcal{R}_n(B)$ .

# Радемахеровская сложность

## Радемахеровское среднее

Следующие свойства непосредственно вытекают из приведённого определения.

### Утверждение 5.7.

Пусть  $n \in \mathbb{N}$ ,  $A, B \subset \mathbb{R}^n$  – ограниченные множества и  $c \in \mathbb{R}$ .

Тогда

- ❶  $\mathcal{R}_n(cA) = |c|\mathcal{R}_n(A)$ ;
- ❷ если  $A \subseteq B$ , то  $\mathcal{R}_n(A) \leq \mathcal{R}_n(B)$ ;
- ❸  $\mathcal{R}_n(A + B) = \mathcal{R}_n(A) + \mathcal{R}_n(B)$ .

Установим связь между радемахеровскими средними ограниченного множества и его выпуклой оболочкой.

# Радемахеровская сложность

Радемахеровское среднее

## Утверждение 5.8.

Пусть  $n \in \mathbb{N}$ ,  $A \subset \mathbb{R}^n$  – ограниченное множество. Тогда

$$\mathcal{R}_n(A) = \mathcal{R}_n(\text{conv}A).$$

# Радемахеровская сложность

## Радемахеровское среднее

### Утверждение 5.8.

Пусть  $n \in \mathbb{N}$ ,  $A \subset \mathbb{R}^n$  – ограниченное множество. Тогда

$$\mathcal{R}_n(A) = \mathcal{R}_n(\text{conv}A).$$

◀ Напомним, что под выпуклой оболочкой множества  $A$  понимается множество

$$\text{conv}A := \{ \lambda \mathbf{a} + (1 - \lambda) \mathbf{a}' : \mathbf{a}, \mathbf{a}' \in A, \lambda \in [0, 1] \}.$$

# Радемахеровская сложность

## Радемахеровское среднее

### Утверждение 5.8.

Пусть  $n \in \mathbb{N}$ ,  $A \subset \mathbb{R}^n$  – ограниченное множество. Тогда

$$\mathcal{R}_n(A) = \mathcal{R}_n(\text{conv}A).$$

◀ Напомним, что под выпуклой оболочкой множества  $A$  понимается множество

$$\text{conv}A := \{ \lambda \mathbf{a} + (1 - \lambda) \mathbf{a}' : \mathbf{a}, \mathbf{a}' \in A, \lambda \in [0, 1] \}.$$

Для доказательства утверждения достаточно заметить, что

# Радемахеровская сложность

## Радемахеровское среднее

$$\begin{aligned}\sup_{\mathbf{a}, \mathbf{a}' \in A, \lambda \in [0,1]} \langle \lambda \mathbf{a} + (1 - \lambda) \mathbf{a}', \epsilon \rangle &= \sup_{\mathbf{a}, \mathbf{a}' \in A, \lambda \in [0,1]} \lambda \langle \mathbf{a}, \epsilon \rangle + (1 - \lambda) \langle \mathbf{a}', \epsilon \rangle \\ &= \sup_{\mathbf{a}, \mathbf{a}' \in A} \max_{\lambda \in [0,1]} \lambda \langle \mathbf{a}, \epsilon \rangle + (1 - \lambda) \langle \mathbf{a}', \epsilon \rangle \\ &= \sup_{\mathbf{a}, \mathbf{a}' \in A} \max_{\lambda \in \{0,1\}} \lambda \langle \mathbf{a}, \epsilon \rangle + (1 - \lambda) \langle \mathbf{a}', \epsilon \rangle \\ &= \sup_{\mathbf{a} \in A} \langle \mathbf{a}, \epsilon \rangle,\end{aligned}$$

где через  $\epsilon$  обозначен вектор, состоящий из  $n$  независимых радемахеровских случайных величин.



# Радемахеровская сложность

## Радемахеровское среднее

Существует альтернативное определение, в котором под радемахеровским средним понимается число

$$\widehat{\mathcal{R}}_n(A) := \mathbf{E} \left[ \sup_{(a_1, \dots, a_n) \in A} \frac{1}{n} \left| \sum_{i=1}^n \varepsilon_i a_i \right| \right].$$

# Радемахеровская сложность

## Радемахеровское среднее

Существует альтернативное определение, в котором под радемахеровским средним понимается число

$$\widehat{\mathcal{R}}_n(A) := \mathbf{E} \left[ \sup_{(a_1, \dots, a_n) \in A} \frac{1}{n} \left| \sum_{i=1}^n \varepsilon_i a_i \right| \right].$$

Следующее утверждение устанавливает связь между этими двумя определениями.



# Радемахеровская сложность

## Радемахеровское среднее

Существует альтернативное определение, в котором под радемахеровским средним понимается число

$$\widehat{\mathcal{R}}_n(A) := \mathbf{E} \left[ \sup_{(a_1, \dots, a_n) \in A} \frac{1}{n} \left| \sum_{i=1}^n \varepsilon_i a_i \right| \right].$$

Следующее утверждение устанавливает связь между этими двумя определениями.

### Утверждение 5.9.

Пусть  $n \in \mathbb{N}$  и  $A \subset \mathbb{R}^n$  – ограниченное множество. Тогда

$$\mathcal{R}_n(A) \leq \mathcal{R}_n(A \cup -A) = \widehat{\mathcal{R}}_n(A).$$

# Радемахеровская сложность

## Радемахеровское среднее

Отмеченные свойства радемахеровского среднего подсказывают идею его приближённого вычисления.

# Радемахеровская сложность

## Радемахеровское среднее

Отмеченные свойства радемахеровского среднего подсказывают идею его приближённого вычисления.

Вначале можно попытаться представить ограниченное множество в виде набора несвязных компонент и свести задачу к вычислению радемахеровского среднего для каждой из них.

# Радемахеровская сложность

## Радемахеровское среднее

Отмеченные свойства радемахеровского среднего подсказывают идею его приближённого вычисления.

Вначале можно попытаться представить ограниченное множество в виде набора несвязных компонент и свести задачу к вычислению радемахеровского среднего для каждой из них.

Отдельные компоненты можно попытаться приблизить выпуклыми множествами.

# Радемахеровская сложность

## Радемахеровское среднее

Отмеченные свойства радемахеровского среднего подсказывают идею его приближённого вычисления.

Вначале можно попытаться представить ограниченное множество в виде набора несвязных компонент и свести задачу к вычислению радемахеровского среднего для каждой из них.

Отдельные компоненты можно попытаться приблизить выпуклыми множествами.

Вычисление радемахеровского среднего выпуклого множества очевидно сводится к вычислению радемахеровского среднего конечного множества.

# Радемахеровская сложность

## Радемахеровское среднее

Отмеченные свойства радемахеровского среднего подсказывают идею его приближённого вычисления.

Вначале можно попытаться представить ограниченное множество в виде набора несвязных компонент и свести задачу к вычислению радемахеровского среднего для каждой из них.

Отдельные компоненты можно попытаться приблизить выпуклыми множествами.

Вычисление радемахеровского среднего выпуклого множества очевидно сводится к вычислению радемахеровского среднего конечного множества.

В этой связи особую роль и важность приобретает следующая лемма.

# Радемахеровская сложность

## Радемахеровское среднее

### Лемма 5.4 (лемма о конечном классе, Массар).

Пусть  $A = \{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(N)}\} \subset \mathbb{R}^n$  ( $n, N \in \mathbb{N}$ ) и  $L \in \mathbb{R}_+$  такие, что  $\|\mathbf{a}^{(j)}\|_2 \leq L$  ( $j = 1, \dots, N$ ).

# Радемахеровская сложность

Радемахеровское среднее

## Лемма 5.4 (лемма о конечном классе, Массар).

Пусть  $A = \{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(N)}\} \subset \mathbb{R}^n$  ( $n, N \in \mathbb{N}$ ) и  $L \in \mathbb{R}_+$  такие, что  $\|\mathbf{a}^{(j)}\|_2 \leq L$  ( $j = 1, \dots, N$ ).

Тогда

$$\mathcal{R}_n(A) \leq \frac{L\sqrt{2\ln N}}{n}. \quad (2)$$



# Радемахеровская сложность

## Радемахеровское среднее

### Лемма 5.4 (лемма о конечном классе, Массар).

Пусть  $A = \{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(N)}\} \subset \mathbb{R}^n$  ( $n, N \in \mathbb{N}$ ) и  $L \in \mathbb{R}_+$  такие, что  $\|\mathbf{a}^{(j)}\|_2 \leq L$  ( $j = 1, \dots, N$ ).

Тогда

$$\mathcal{R}_n(A) \leq \frac{L\sqrt{2\ln N}}{n}. \quad (2)$$

◀ Определим случайные величины

$$\xi_j := \frac{1}{n} \sum_{i=1}^n \varepsilon_i a_i^{(j)} \quad (j = 1, \dots, N).$$

# Радемахеровская сложность

## Радемахеровское среднее

### Напоминание

**Лемма(Хёфдинг).** Пусть  $\xi$  – случайная величина и  $\xi \in [a, b]$  (п.н.) для некоторых  $a, b \in \mathbb{R}$ ,  $a < b$ . Тогда для любого  $\varepsilon > 0$  выполняется неравенство

$$\mathbf{E}[e^{\varepsilon(\xi - \mathbf{E}\xi)}] \leq e^{\frac{\varepsilon^2(b-a)^2}{8}}.$$

**Определение.** Случайная величина  $\xi$  называется *субгауссовской с параметром масштабирования  $s > 0$* , если для любого  $\varepsilon > 0$  выполняется неравенство

$$\mathbf{E}[e^{\varepsilon(\xi - \mathbf{E}\xi)}] \leq e^{\frac{\varepsilon^2 s^2}{2}}.$$

**Лемма(максимальное неравенство).** Пусть  $\xi_1, \dots, \xi_n$  – субгауссовские случайные величины с параметром масштабирования  $s > 0$  и нулевым математическим ожиданием.

Тогда

$$\mathbf{E}\left[\max_{1 \leq i \leq n} \xi_i\right] \leq s\sqrt{2 \ln n}.$$

# Радемахеровская сложность

## Радемахеровское среднее

Из леммы Хёффдинга следует, что случайная величина  $\frac{a_i^{(j)}}{n} \varepsilon_i$  является субгауссовской с параметром масштабирования равным  $\frac{|a_i^{(j)}|}{n}$ .

# Радемахеровская сложность

## Радемахеровское среднее

Из леммы Хёффдинга следует, что случайная величина  $\frac{a_i^{(j)}}{n} \varepsilon_i$  является субгауссовской с параметром масштабирования равным  $\frac{|a_i^{(j)}|}{n}$ .

Случайная величина  $\xi_j$  является суммой независимых субгауссовских случайных величин, поэтому, она сама является субгауссовской с параметром масштабирования  $s$  таким, что

$$s^2 = \frac{1}{n^2} \sum_{i=1}^n |a_i^{(j)}|^2 = \frac{\|\mathbf{a}^{(j)}\|_2^2}{n^2} \leq \left(\frac{L}{n}\right)^2.$$

# Радемахеровская сложность

## Радемахеровское среднее

Из леммы Хёффдинга следует, что случайная величина  $\frac{a_i^{(j)}}{n} \varepsilon_i$  является субгауссовской с параметром масштабирования равным  $\frac{|a_i^{(j)}|}{n}$ .

Случайная величина  $\xi_j$  является суммой независимых субгауссовских случайных величин, поэтому, она сама является субгауссовской с параметром масштабирования  $s$  таким, что

$$s^2 = \frac{1}{n^2} \sum_{i=1}^n |a_i^{(j)}|^2 = \frac{\|\mathbf{a}^{(j)}\|_2^2}{n^2} \leq \left(\frac{L}{n}\right)^2.$$

Используя представление  $\mathcal{R}_n(\mathbf{A}) = \max\{\xi_1, \dots, \xi_N\}$  и применяя максимальное неравенство (лемма 2.2), получим неравенство (2).



# Радемахеровская сложность

## Радемахеровское среднее

Аналогичное утверждение может быть получено и для альтернативного определения радемахеровского среднего.

# Радемахеровская сложность

## Радемахеровское среднее

Аналогичное утверждение может быть получено и для альтернативного определения радемахеровского среднего.

### Следствие 5.1.

Пусть  $A = \{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(N)}\} \subset \mathbb{R}^n$  ( $n, N \in \mathbb{N}$ ) и  $L \in \mathbb{R}_+$  такие, что  $\|\mathbf{a}^{(j)}\|_2 \leq L$  ( $j = 1, \dots, N$ ).

# Радемахеровская сложность

## Радемахеровское среднее

Аналогичное утверждение может быть получено и для альтернативного определения радемахеровского среднего.

### Следствие 5.1.

Пусть  $A = \{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(N)}\} \subset \mathbb{R}^n$  ( $n, N \in \mathbb{N}$ ) и  $L \in \mathbb{R}_+$  такие, что  $\|\mathbf{a}^{(j)}\|_2 \leq L$  ( $j = 1, \dots, N$ ).

Тогда

$$\hat{\mathcal{R}}_n(A) \leq \frac{L\sqrt{2\ln(2N)}}{n}, \quad (3)$$



# Радемахеровская сложность

## Радемахеровское среднее

Аналогичное утверждение может быть получено и для альтернативного определения радемахеровского среднего.

### Следствие 5.1.

Пусть  $A = \{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(N)}\} \subset \mathbb{R}^n$  ( $n, N \in \mathbb{N}$ ) и  $L \in \mathbb{R}_+$  такие, что  $\|\mathbf{a}^{(j)}\|_2 \leq L$  ( $j = 1, \dots, N$ ).

Тогда

$$\widehat{\mathcal{R}}_n(A) \leq \frac{L\sqrt{2\ln(2N)}}{n}, \quad (3)$$

в частности

$$\widehat{\mathcal{R}}_n(A) \leq \frac{2L\sqrt{\ln N}}{n} \quad (N \geq 2). \quad (4)$$

# Радемахеровская сложность

## Радемахеровское среднее

◀ Подставляя  $\hat{\mathcal{R}}_n(A) = \mathcal{R}_n(A \cup -A)$  в неравенство (2), получим первое неравенство (3).

# Радемахеровская сложность

## Радемахеровское среднее

◀ Подставляя  $\hat{\mathcal{R}}_n(A) = \mathcal{R}_n(A \cup -A)$  в неравенство (2), получим первое неравенство (3).

Для доказательства второго неравенства (4) достаточно заметить, что  $\ln(2N) \leq 2 \ln N$  при  $N \geq 2$ .



# Содержание

- 1 Радемахеровская сложность
  - Радемахеровское среднее
  - **Радемахеровская сложность**
- 2 Эмпирические и предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности

# Радемахеровская сложность

## Радемахеровская сложность

Радемахеровская сложность определяется для класса функций, которые должны быть ограничены в совокупности.

# Радемахеровская сложность

## Радемахеровская сложность

Радемахеровская сложность определяется для класса функций, которые должны быть ограничены в совокупности.

Формально говоря, в определении не требуется наличие класса гипотез и функции потерь, которые порождают этот функциональный класс.

# Радемахеровская сложность

## Радемахеровская сложность

Радемахеровская сложность определяется для класса функций, которые должны быть ограничены в совокупности.

Формально говоря, в определении не требуется наличие класса гипотез и функции потерь, которые порождают этот функциональный класс.

Однако подобная интерпретация является полезной.

# Радемахеровская сложность

## Радемахеровская сложность

### Определение 5.8.

Пусть  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ) и  $n \in \mathbb{N}$ .



# Радемахеровская сложность

## Радемахеровская сложность

### Определение 5.8.

Пусть  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ) и  $n \in \mathbb{N}$ .

**Эмпирической радемахеровской сложностью** класса функций  $\mathcal{F}$  называется функция

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) := \mathcal{R}_n(\{(f(z_1), \dots, f(z_n)) : f \in \mathcal{F}\}) \quad (\mathbf{z} \in Z^n).$$

# Радемахеровская сложность

## Радемахеровская сложность

### Определение 5.8.

Пусть  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ) и  $n \in \mathbb{N}$ .

**Эмпирической радемахеровской сложностью** класса функций  $\mathcal{F}$  называется функция

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) := \mathcal{R}_n(\{(f(z_1), \dots, f(z_n)) : f \in \mathcal{F}\}) \quad (\mathbf{z} \in Z^n).$$

**Радемахеровской сложностью** класса функций  $\mathcal{F}$  относительно вероятностной меры  $P \in \mathcal{M}_+^1(Z)$  называется число

$$\mathcal{R}_n(\mathcal{F}, P) := \mathbf{E}_{\mathbf{z} \sim P^n} \left[ \mathcal{R}_n(\mathcal{F}, \mathbf{z}) \right].$$

# Радемахеровская сложность

## Радемахеровская сложность

Если из контекста понятно, о какой вероятностной мере  $P$  идёт речь, то для радемахеровской сложности будет использоваться сокращённое обозначение  $\mathcal{R}_n(\mathcal{F})$ .

# Радемахеровская сложность

## Радемахеровская сложность

### Утверждение 5.10.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ),  $\varepsilon > 0$  и  $n \in \mathbb{N}$ .

# Радемахеровская сложность

## Радемахеровская сложность

### Утверждение 5.10.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ),  $\varepsilon > 0$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \mathcal{R}_n(\mathcal{F}) - \mathcal{R}_n(\mathcal{F}, \mathbf{z}) \geq \varepsilon \right\} \leq \exp \left( \frac{-2n\varepsilon^2}{(b-a)^2} \right). \quad (5)$$

# Радемахеровская сложность

## Радемахеровская сложность

### Утверждение 5.10.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ),  $\varepsilon > 0$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \mathcal{R}_n(\mathcal{F}) - \mathcal{R}_n(\mathcal{F}, \mathbf{z}) \geq \varepsilon \right\} \leq \exp \left( \frac{-2n\varepsilon^2}{(b-a)^2} \right). \quad (5)$$

◀ Рассмотрим произвольные  $\mathbf{z}, \mathbf{z}' \in Z^n$ , которые отличаются только в  $k$ -ой ( $1 \leq k \leq n$ ) компоненте.

# Радемахеровская сложность

## Радемахеровская сложность

Тогда

$$\begin{aligned}\mathcal{R}_n(\mathcal{F}, \mathbf{z}) &= \frac{1}{n} \mathbf{E}_{\epsilon \sim Q_n} \left[ \sup_{f \in \mathcal{F}} \sum_{i=1}^n \epsilon_i (f(\mathbf{z}'_i) + f(\mathbf{z}_i) - f(\mathbf{z}'_i)) \right] \\ &\leq \frac{1}{n} \mathbf{E}_{\epsilon \sim Q_n} \left[ \sup_{f \in \mathcal{F}} \sum_{i=1}^n \epsilon_i f(\mathbf{z}'_i) \right] + \frac{1}{n} \mathbf{E}_{\epsilon \sim Q_n} \left[ \sup_{f \in \mathcal{F}} \sum_{i=1}^n \epsilon_i (f(\mathbf{z}_i) - f(\mathbf{z}'_i)) \right] \\ &\leq \mathcal{R}_n(\mathcal{F}, \mathbf{z}') + \frac{1}{n} \mathbf{E}_{\epsilon \sim Q_n} \left[ \sup_{f \in \mathcal{F}} \sum_{k=1}^n \epsilon_k (f(\mathbf{z}_k) - f(\mathbf{z}'_k)) \right] \\ &\leq \mathcal{R}_n(\mathcal{F}, \mathbf{z}') + \frac{|b - a|}{n}.\end{aligned}$$

# Радемахеровская сложность

## Радемахеровская сложность

Таким образом, функция  $\hat{\mathcal{R}}_n(\mathcal{F}, \cdot)$  обладает свойством покоординатных приращений с коэффициентами  $c_i = \frac{|b-a|}{n}$  ( $i = 1, \dots, n$ ).



# Радемахеровская сложность

## Радемахеровская сложность

Таким образом, функция  $\hat{\mathcal{R}}_n(\mathcal{F}, \cdot)$  обладает свойством покоординатных приращений с коэффициентами  $c_i = \frac{|b-a|}{n}$  ( $i = 1, \dots, n$ ).

Рассмотрим вероятностное пространство  $(Z^n, \otimes^n \mathcal{Z}, P^n)$ .

# Радемахеровская сложность

## Радемахеровская сложность

Таким образом, функция  $\hat{\mathcal{R}}_n(\mathcal{F}, \cdot)$  обладает свойством покоординатных приращений с коэффициентами  $c_i = \frac{|b-a|}{n}$  ( $i = 1, \dots, n$ ).

Рассмотрим вероятностное пространство  $(Z^n, \otimes^n \mathcal{Z}, P^n)$ .

Для функции  $\mathcal{R}_n(\mathcal{F}, \cdot)$  и функций координатных проекций запишем неравенство МакДиармида.

# Радемахеровская сложность

## Радемахеровская сложность

Таким образом, функция  $\hat{\mathcal{R}}_n(\mathcal{F}, \cdot)$  обладает свойством покоординатных приращений с коэффициентами  $c_i = \frac{|b-a|}{n}$  ( $i = 1, \dots, n$ ).

Рассмотрим вероятностное пространство  $(Z^n, \otimes^n \mathcal{Z}, P^n)$ .

Для функции  $\mathcal{R}_n(\mathcal{F}, \cdot)$  и функций координатных проекций запишем неравенство МакДиармида.

Получим искомое неравенство (5).



# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.5.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\mathcal{H} \simeq_{l_{01}} \mathcal{F}$  и  $n \in \mathbb{N}$ .

# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.5.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\mathcal{H} \simeq_{l_01} \mathcal{F}$  и  $n \in \mathbb{N}$ .

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) = \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \quad (\mathbf{z} \in Z^n, \mathbf{x} = \mathbf{z}|_X). \quad (6)$$

# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.5.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\mathcal{H} \simeq_{l_01} \mathcal{F}$  и  $n \in \mathbb{N}$ .

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) = \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \quad (\mathbf{z} \in Z^n, \mathbf{x} = \mathbf{z}|_X). \quad (6)$$

Как следствие,

$$\mathcal{R}_n(\mathcal{F}, P) = \mathcal{R}_n(\mathcal{H}, P_X). \quad (7)$$

# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.5.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\mathcal{H} \simeq_{l_{01}} \mathcal{F}$  и  $n \in \mathbb{N}$ .

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) = \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \quad (\mathbf{z} \in Z^n, \mathbf{x} = \mathbf{z}|_X). \quad (6)$$

Как следствие,

$$\mathcal{R}_n(\mathcal{F}, P) = \mathcal{R}_n(\mathcal{H}, P_X). \quad (7)$$

◀ Зафиксируем произвольный набор  $\mathbf{z} \in Z^n$  и заметим, что  $1 - 2y_i \in \{\pm 1\}$  ( $i = 1, \dots, n$ ).

# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.5.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\mathcal{H} \simeq_{l_{01}} \mathcal{F}$  и  $n \in \mathbb{N}$ .

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) = \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \quad (\mathbf{z} \in Z^n, \mathbf{x} = \mathbf{z}|_X). \quad (6)$$

Как следствие,

$$\mathcal{R}_n(\mathcal{F}, P) = \mathcal{R}_n(\mathcal{H}, P_X). \quad (7)$$

◀ Зафиксируем произвольный набор  $\mathbf{z} \in Z^n$  и заметим, что  $1 - 2y_i \in \{\pm 1\}$  ( $i = 1, \dots, n$ ).

Следовательно, если  $\varepsilon_1, \dots, \varepsilon_n$  – независимые радемахеровские случайные величины,



# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.5.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\mathcal{H} \simeq_{l_{01}} \mathcal{F}$  и  $n \in \mathbb{N}$ .

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) = \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \quad (\mathbf{z} \in Z^n, \mathbf{x} = \mathbf{z}|_X). \quad (6)$$

Как следствие,

$$\mathcal{R}_n(\mathcal{F}, P) = \mathcal{R}_n(\mathcal{H}, P_X). \quad (7)$$

◀ Зафиксируем произвольный набор  $\mathbf{z} \in Z^n$  и заметим, что  $1 - 2y_i \in \{\pm 1\}$  ( $i = 1, \dots, n$ ).

Следовательно, если  $\varepsilon_1, \dots, \varepsilon_n$  – независимые радемахеровские случайные величины,

то и случайные величины  $(1 - 2y_1)\varepsilon_1, \dots, (1 - 2y_n)\varepsilon_n$  также будут независимыми и радемахеровскими.

# Радемахеровская сложность

## Радемахеровская сложность

Для любых  $y, y' \in \{0, 1\}$  имеет место равенство

$$I_{01}(y, y') = y + y' - 2yy',$$

# Радемахеровская сложность

## Радемахеровская сложность

Для любых  $y, y' \in \{0, 1\}$  имеет место равенство

$$l_{01}(y, y') = y + y' - 2yy',$$

а значит

$$\begin{aligned}\mathcal{R}_n(\mathcal{F}, \mathbf{z}) &= \mathbf{E} \left[ \sup_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \varepsilon_i l_{01}(y_i, h(x_i)) \right] \\&= \mathbf{E} \left[ \sup_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \varepsilon_i (y_i + (1 - 2y_i)h(x_i)) \right] \\&= \mathbf{E} \left[ \frac{1}{n} \sum_{i=1}^n y_i \varepsilon_i \right] + \mathbf{E} \left[ \sup_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \varepsilon_i (1 - 2y_i)h(x_i) \right] \\&= 0 + \mathbf{E} \left[ \sup_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \varepsilon_i h(x_i) \right] \\&= \mathcal{R}_n(\mathcal{H}, \mathbf{x}).\end{aligned}$$

# Радемахеровская сложность

## Радемахеровская сложность

Перейдём к доказательству равенства (7).

# Радемахеровская сложность

## Радемахеровская сложность

Перейдём к доказательству равенства (7).

Запишем

$$\begin{aligned}\mathcal{R}_n(\mathcal{F}, P) &= \left| \text{(6)} \right| = \mathbf{E}_{\mathbf{z} \sim P^n} \left[ \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \right] = \left| \begin{array}{c} \text{Теорема} \\ \text{Фубини} \end{array} \right| \\ &= \mathbf{E}_{(x_n, y_n) \sim P} \cdots \mathbf{E}_{(x_1, y_1) \sim P} \left[ \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \right] = \left| \begin{array}{c} \text{Теорема о} \\ \text{дезинтеграции} \end{array} \right| \\ &= \mathbf{E}_{(x_n, y_n) \sim P} \cdots \mathbf{E}_{(x_2, y_2) \sim P} \mathbf{E}_{x_1 \sim P_X} \mathbf{E}_{y_1 \sim P_{Y|X}(x_1, \cdot)} \left[ \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \right]\end{aligned}$$

# Радемахеровская сложность

## Радемахеровская сложность

$$\begin{aligned}\mathcal{R}_n(\mathcal{F}, \mathbf{P}) &= \left| \begin{array}{c} \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \\ \text{не зависит от } y_1 \end{array} \right| = \mathbf{E}_{(x_n, y_n) \sim \mathbf{P}} \cdots \mathbf{E}_{(x_2, y_2) \sim \mathbf{P}} \mathbf{E}_{x_1 \sim \mathbf{P}_X} \left[ \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \right] \\ &= \left| \begin{array}{c} \text{Теорема} \\ \text{Фубини} \end{array} \right| = \mathbf{E}_{x_1 \sim \mathbf{P}_X} \mathbf{E}_{(x_n, y_n) \sim \mathbf{P}} \cdots \mathbf{E}_{(x_2, y_2) \sim \mathbf{P}} \left[ \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \right] \\ &= \dots \\ &= \mathbf{E}_{x_1 \sim \mathbf{P}_X} \cdots \mathbf{E}_{x_n \sim \mathbf{P}_X} \left[ \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \right] \\ &= \left| \begin{array}{c} \text{Теорема} \\ \text{Фубини} \end{array} \right| = \mathbf{E}_{\mathbf{x} \sim \mathbf{P}_X^n} \left[ \mathcal{R}_n(\mathcal{H}, \mathbf{x}) \right] \\ &= \mathcal{R}_n(\mathcal{H}, \mathbf{P}_X).\end{aligned}$$



# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.6.

Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^{\mathbb{Z}}$ .

# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.6.

Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^{\mathbb{Z}}$ .

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \sqrt{\frac{2 \ln(\Gamma_{\mathcal{F}}(n))}{n}} \quad (\mathbf{z} \in \mathbb{Z}^n). \quad (8)$$



# Радемахеровская сложность

## Радемахеровская сложность

### Лемма 5.6.

Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^{\mathbb{Z}}$ .

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \sqrt{\frac{2 \ln(\Gamma_{\mathcal{F}}(n))}{n}} \quad (\mathbf{z} \in \mathbb{Z}^n). \quad (8)$$

Кроме того, если дополнительно  $\text{vc}(\mathcal{F}) < \infty$ , то

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \sqrt{\frac{2 \text{vc}(\mathcal{F}) \ln(n+1)}{n}} \quad (\mathbf{z} \in \mathbb{Z}^n). \quad (9)$$

# Радемахеровская сложность

## Радемахеровская сложность

◀ Неравенство (8) вытекает из леммы Массара о конечном классе.

# Радемахеровская сложность

## Радемахеровская сложность

◀ Неравенство (8) вытекает из леммы Массара о конечном классе.

В качестве  $A$  выступает  $\mathcal{F}(\mathbf{z})$ , в качестве  $L$  выступает  $\sqrt{n}$ , и  $N = |\mathcal{F}(\mathbf{z})| \leq \Gamma_{\mathcal{F}}(n)$ .

# Радемахеровская сложность

## Радемахеровская сложность

◀ Неравенство (8) вытекает из леммы Массара о конечном классе.

В качестве  $A$  выступает  $\mathcal{F}(\mathbf{z})$ , в качестве  $L$  выступает  $\sqrt{n}$ , и  $N = |\mathcal{F}(\mathbf{z})| \leq \Gamma_{\mathcal{F}}(n)$ .

Неравенство (9) следует из предыдущего неравенства (8) и леммы Сауэра-Шелаха.



# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
  - Неравенство симметризации
  - Эмпирические оценки
  - Предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности

# Эмпирические и предсказывающие оценки

Отталкиваясь от понятия радемахеровской сложности, перейдём к получению эмпирических и предсказывающих оценок, о которых говорилось в начале этой лекции.

# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
  - Неравенство симметризации
  - Эмпирические оценки
  - Предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

### Теорема 5.6.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$  и  $n \in \mathbb{N}$ .



# Эмпирические и предсказывающие оценки

## Неравенство симметризации

### Теорема 5.6.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$  и  $n \in \mathbb{N}$ .

Определим две функции

$$\Delta_{\mathcal{F}}^+(\mathbf{z}) := \sup_{f \in \mathcal{F}} (R(f) - r(f, \mathbf{z})), \quad \Delta_{\mathcal{F}}^-(\mathbf{z}) := \sup_{f \in \mathcal{F}} (r(f, \mathbf{z}) - R(f)) \quad (\mathbf{z} \in Z^n).$$

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

### Теорема 5.6.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$  и  $n \in \mathbb{N}$ .

Определим две функции

$$\Delta_{\mathcal{F}}^+(\mathbf{z}) := \sup_{f \in \mathcal{F}} (R(f) - r(f, \mathbf{z})), \quad \Delta_{\mathcal{F}}^-(\mathbf{z}) := \sup_{f \in \mathcal{F}} (r(f, \mathbf{z}) - R(f)) \quad (\mathbf{z} \in Z^n).$$

Тогда

$$\mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] \leq 2 \mathcal{R}_n(\mathcal{F}), \quad (10)$$

$$\mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^-(\mathbf{z})] \leq 2 \mathcal{R}_n(\mathcal{F}). \quad (11)$$

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

### Теорема 5.6.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$  и  $n \in \mathbb{N}$ .

Определим две функции

$$\Delta_{\mathcal{F}}^+(\mathbf{z}) := \sup_{f \in \mathcal{F}} (R(f) - r(f, \mathbf{z})), \quad \Delta_{\mathcal{F}}^-(\mathbf{z}) := \sup_{f \in \mathcal{F}} (r(f, \mathbf{z}) - R(f)) \quad (\mathbf{z} \in Z^n).$$

Тогда

$$\mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] \leq 2 \mathcal{R}_n(\mathcal{F}), \quad (10)$$

$$\mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^-(\mathbf{z})] \leq 2 \mathcal{R}_n(\mathcal{F}). \quad (11)$$

◀ Докажем неравенство (10). Неравенство (11) проверяется аналогичным образом.

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Будем использовать вспомогательное вероятностное пространство  $(Z^{2n}, \otimes^{2n} \mathcal{Z}, \mathbf{P}^{2n})$ .

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Будем использовать вспомогательное вероятностное пространство  $(Z^{2n}, \otimes^{2n} \mathcal{Z}, P^{2n})$ .

Его элементарные события имеют вид  $(z_1, \dots, z_n, z'_1, \dots, z'_n)$ .

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Будем использовать вспомогательное вероятностное пространство  $(Z^{2n}, \otimes^{2n} \mathcal{Z}, P^{2n})$ .

Его элементарные события имеют вид  $(z_1, \dots, z_n, z'_1, \dots, z'_n)$ .

Соответствующие функции координатных проекций  $Z_1, \dots, Z_n, Z'_1, \dots, Z'_n$  являются независимыми случайными элементами с распределением вероятностей  $P$ .

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Будем использовать вспомогательное вероятностное пространство  $(Z^{2n}, \otimes^{2n} \mathcal{Z}, P^{2n})$ .

Его элементарные события имеют вид  $(z_1, \dots, z_n, z'_1, \dots, z'_n)$ .

Соответствующие функции координатных проекций  $Z_1, \dots, Z_n, Z'_1, \dots, Z'_n$  являются независимыми случайными элементами с распределением вероятностей  $P$ .

Случайные элементы  $\mathbf{Z} := (Z_1, \dots, Z_n)$  и  $\mathbf{Z}' := (Z'_1, \dots, Z'_n)$  также являются независимыми и одинаково распределёнными.

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Используя свойства условного математического ожидания, запишем



# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Используя свойства условного математического ожидания, запишем

$$\begin{aligned} \mathbf{E} \left[ \sup_{f \in \mathcal{F}} (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \mid \mathbf{Z} \right] &\geq \left| \begin{array}{c} \text{УТВ. 1.17} \\ (3) \end{array} \right| \geq \\ \sup_{f \in \mathcal{F}} \mathbf{E} \left[ (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \mid \mathbf{Z} \right] &= \left| \begin{array}{c} \text{УТВ. 1.17} \\ (1, 4) \end{array} \right| = \\ \sup_{f \in \mathcal{F}} (R(f) - r(f, \mathbf{Z})) &= \\ \Delta_{\mathcal{F}}^+(\mathbf{Z}) &\quad (\text{п.н.}). \end{aligned} \tag{12}$$

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Используя свойства условного математического ожидания, запишем

$$\begin{aligned} \mathbf{E} \left[ \sup_{f \in \mathcal{F}} (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \mid \mathbf{Z} \right] &\geq \left| \begin{array}{c} \text{утв. 1.17} \\ (3) \end{array} \right| \geq \\ \sup_{f \in \mathcal{F}} \mathbf{E} \left[ (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \mid \mathbf{Z} \right] &= \left| \begin{array}{c} \text{утв. 1.17} \\ (1, 4) \end{array} \right| = \\ \sup_{f \in \mathcal{F}} (R(f) - r(f, \mathbf{Z})) &= \\ \Delta_{\mathcal{F}}^+(\mathbf{Z}) &\quad (\text{п.н.}). \end{aligned} \tag{12}$$

Переходя к математическому ожиданию в левой и правой части неравенства (12), получим

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Используя свойства условного математического ожидания, запишем

$$\begin{aligned} \mathbf{E} \left[ \sup_{f \in \mathcal{F}} (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \mid \mathbf{Z} \right] &\geq \left| \begin{array}{c} \text{утв. 1.17} \\ (3) \end{array} \right| \geq \\ \sup_{f \in \mathcal{F}} \mathbf{E} \left[ (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \mid \mathbf{Z} \right] &= \left| \begin{array}{c} \text{утв. 1.17} \\ (1, 4) \end{array} \right| = \\ \sup_{f \in \mathcal{F}} (R(f) - r(f, \mathbf{Z})) &= \\ \Delta_{\mathcal{F}}^+(\mathbf{Z}) &\quad (\text{п.н.}). \end{aligned} \tag{12}$$

Переходя к математическому ожиданию в левой и правой части неравенства (12), получим

$$\mathbf{E} \left[ \Delta_{\mathcal{F}}^+(\mathbf{Z}) \right] \leq \mathbf{E} \left[ \sup_{f \in \mathcal{F}} (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \right]. \tag{13}$$

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Для любых  $\epsilon_1, \dots, \epsilon_n \in \{\pm 1\}$  и любой функции  $f \in \mathcal{F}$  случайные величины

$$\epsilon_1(f(Z_1) - f(Z'_1)), \dots, \epsilon_n(f(Z_n) - f(Z'_n))$$

будут независимыми и одинаково распределёнными.

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Для любых  $\epsilon_1, \dots, \epsilon_n \in \{\pm 1\}$  и любой функции  $f \in \mathcal{F}$  случайные величины

$$\epsilon_1(f(Z_1) - f(Z'_1)), \dots, \epsilon_n(f(Z_n) - f(Z'_n))$$

будут независимыми и одинаково распределёнными.

Следовательно, одинаково распределены и случайные величины

$$\frac{1}{n} \sum_{i=1}^n (f(Z'_i) - f(Z_i)) \sim \frac{1}{n} \sum_{i=1}^n \epsilon_i (f(Z'_i) - f(Z_i)).$$

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Для любых  $\epsilon_1, \dots, \epsilon_n \in \{\pm 1\}$  и любой функции  $f \in \mathcal{F}$  случайные величины

$$\epsilon_1(f(Z_1) - f(Z'_1)), \dots, \epsilon_n(f(Z_n) - f(Z'_n))$$

будут независимыми и одинаково распределёнными.

Следовательно, одинаково распределены и случайные величины

$$\frac{1}{n} \sum_{i=1}^n (f(Z'_i) - f(Z_i)) \sim \frac{1}{n} \sum_{i=1}^n \epsilon_i (f(Z'_i) - f(Z_i)).$$

Далее, получим

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

$$\begin{aligned} \mathbf{E} \left[ \sup_{f \in \mathcal{F}} (r(f, \mathbf{Z}') - r(f, \mathbf{Z})) \right] &= \\ \mathbf{E} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n (f(Z'_i) - f(Z_i)) \right] &= \\ \mathbf{E} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i (f(Z'_i) - f(Z_i)) \right] &\leqslant \\ \mathbf{E} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i f(Z'_i) \right] + \mathbf{E} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n -\epsilon_i f(Z_i) \right]. \end{aligned} \tag{14}$$

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Объединим неравенства (13) и (14).



# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Объединим неравенства (13) и (14).

Применяя теорему о замене переменных в интеграле Лебега, запишем

$$\mathbf{E}_{\mathbf{z} \sim \mathcal{P}^n} \left[ \Delta_{\mathcal{F}}^+(\mathbf{z}) \right] \leq \mathbf{E}_{\mathbf{z} \sim \mathcal{P}^n} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i f(\mathbf{z}_i) + \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n -\epsilon_i f(\mathbf{z}_i) \right] \quad (15)$$

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Объединим неравенства (13) и (14).

Применяя теорему о замене переменных в интеграле Лебега, запишем

$$\mathbf{E}_{\mathbf{z} \sim P^n} \left[ \Delta_{\mathcal{F}}^+(\mathbf{z}) \right] \leq \mathbf{E}_{\mathbf{z} \sim P^n} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i f(z_i) + \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n -\epsilon_i f(z_i) \right] \quad (15)$$

Проинтегрируем левую и правую части неравенства (15) по вероятностной мере  $Q_n$ .

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Объединим неравенства (13) и (14).

Применяя теорему о замене переменных в интеграле Лебега, запишем

$$\mathbf{E}_{\mathbf{z} \sim P^n} \left[ \Delta_{\mathcal{F}}^+(\mathbf{z}) \right] \leq \mathbf{E}_{\mathbf{z} \sim P^n} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i f(z_i) + \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n -\epsilon_i f(z_i) \right] \quad (15)$$

Проинтегрируем левую и правую части неравенства (15) по вероятностной мере  $Q_n$ .

В качестве переменной интегрирования возьмём  $\epsilon = (\epsilon_1, \dots, \epsilon_n)$ .

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

Объединим неравенства (13) и (14).

Применяя теорему о замене переменных в интеграле Лебега, запишем

$$\mathbf{E}_{\mathbf{z} \sim P^n} \left[ \Delta_{\mathcal{F}}^+(\mathbf{z}) \right] \leq \mathbf{E}_{\mathbf{z} \sim P^n} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i f(z_i) + \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n -\epsilon_i f(z_i) \right] \quad (15)$$

Проинтегрируем левую и правую части неравенства (15) по вероятностной мере  $Q_n$ .

В качестве переменной интегрирования возьмём  $\epsilon = (\epsilon_1, \dots, \epsilon_n)$ .

Применяя теорему Фубини для замены порядка интегрирования, получим

# Эмпирические и предсказывающие оценки

## Неравенство симметризации

$$\begin{aligned} \mathbf{E}_{\mathbf{z} \sim \mathcal{P}^n} \left[ \Delta_{\mathcal{F}}^+(\mathbf{z}) \right] &\leq \\ \mathbf{E}_{\epsilon \sim \mathcal{Q}_n} \mathbf{E}_{\mathbf{z} \sim \mathcal{P}^n} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i f(\mathbf{z}_i) + \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n -\epsilon_i f(\mathbf{z}_i) \right] &= \\ \mathbf{E}_{\mathbf{z} \sim \mathcal{P}^n} \mathbf{E}_{\epsilon \sim \mathcal{Q}_n} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \epsilon_i f(\mathbf{z}_i) \right] + \mathbf{E}_{\mathbf{z} \sim \mathcal{P}^n} \mathbf{E}_{\epsilon \sim \mathcal{Q}_n} \left[ \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n -\epsilon_i f(\mathbf{z}_i) \right] &= \\ 2 \mathcal{R}_n(\mathcal{F}). \end{aligned} \tag{16}$$



# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
  - Неравенство симметризации
  - Эмпирические оценки
  - Предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Утверждение 5.11.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Утверждение 5.11.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta, \quad (17)$$

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^-(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta. \quad (18)$$



# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Утверждение 5.11.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta, \quad (17)$$

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^-(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta. \quad (18)$$

◀ Установим неравенство (17).

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Утверждение 5.11.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta, \quad (17)$$

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^-(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta. \quad (18)$$

◀ Установим неравенство (17).

Неравенство (18) проверяется аналогичным образом.

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Рассмотрим произвольные  $\mathbf{z}, \mathbf{z}' \in Z^n$ , которые отличаются только в  $i$ -ой ( $1 \leq i \leq n$ ) компоненте.

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Рассмотрим произвольные  $\mathbf{z}, \mathbf{z}' \in Z^n$ , которые отличаются только в  $i$ -ой ( $1 \leq i \leq n$ ) компоненте.

Тогда

$$\begin{aligned}\Delta_{\mathcal{F}}^+(\mathbf{z}) &= \sup_{f \in \mathcal{F}} \left[ (R(f) - r(f, \mathbf{z}')) + (r(f, \mathbf{z}') - r(f, \mathbf{z})) \right] \\ &\leq \Delta_{\mathcal{F}}^+(\mathbf{z}') + \sup_{f \in \mathcal{F}} \frac{1}{n} (f(z_i) - f(z'_i)) \\ &\leq \Delta_{\mathcal{F}}^+(\mathbf{z}') + \frac{1}{n}.\end{aligned}$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Рассмотрим произвольные  $\mathbf{z}, \mathbf{z}' \in Z^n$ , которые отличаются только в  $i$ -ой ( $1 \leq i \leq n$ ) компоненте.

Тогда

$$\begin{aligned}\Delta_{\mathcal{F}}^+(\mathbf{z}) &= \sup_{f \in \mathcal{F}} \left[ (R(f) - r(f, \mathbf{z}')) + (r(f, \mathbf{z}') - r(f, \mathbf{z})) \right] \\ &\leq \Delta_{\mathcal{F}}^+(\mathbf{z}') + \sup_{f \in \mathcal{F}} \frac{1}{n} (f(z_i) - f(z'_i)) \\ &\leq \Delta_{\mathcal{F}}^+(\mathbf{z}') + \frac{1}{n}.\end{aligned}$$

Таким образом, функция  $\Delta_{\mathcal{F}}^+$  обладает свойством покоординатных приращений с коэффициентами  $c_i = \frac{1}{n}$  ( $i = 1, \dots, n$ ).

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Рассмотрим вероятностное пространство  $(Z^n, \otimes^n \mathcal{Z}, P^n)$ .

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Рассмотрим вероятностное пространство  $(Z^n, \otimes^n \mathcal{Z}, P^n)$ .

Для функции  $\Delta_{\mathcal{F}}^+$  и функций координатных проекций запишем неравенство МакДиармида

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) > \mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] + \varepsilon \right\} \leq \exp(-2n\varepsilon^2) \quad (\varepsilon > 0). \quad (19)$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Рассмотрим вероятностное пространство  $(Z^n, \otimes^n \mathcal{Z}, P^n)$ .

Для функции  $\Delta_{\mathcal{F}}^+$  и функций координатных проекций запишем неравенство МакДиармида

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) > \mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] + \varepsilon \right\} \leq \exp(-2n\varepsilon^2) \quad (\varepsilon > 0). \quad (19)$$

Выберем

$$\varepsilon := \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)}, \quad (20)$$



# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Рассмотрим вероятностное пространство  $(Z^n, \otimes^n \mathcal{Z}, P^n)$ .

Для функции  $\Delta_{\mathcal{F}}^+$  и функций координатных проекций запишем неравенство МакДиармида

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) > \mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] + \varepsilon \right\} \leq \exp(-2n\varepsilon^2) \quad (\varepsilon > 0). \quad (19)$$

Выберем

$$\varepsilon := \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)}, \quad (20)$$

тогда

$$-2n\varepsilon^2 = -2n \frac{1}{2n} \ln \left( \frac{1}{\delta} \right) = \ln(\delta). \quad (21)$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Объединяя вместе (19), (20) и (21), получим неравенство

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) \leq \mathbf{E}_{\mathbf{z} \sim P^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta.$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Объединяя вместе (19), (20) и (21), получим неравенство

$$\mathbf{P}^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) \leq \mathbf{E}_{\mathbf{z} \sim \mathbf{P}^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta.$$

Применяя к нему неравенство симметризации, получим требуемое неравенство (17).



# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Теорема 5.7.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Теорема 5.7.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \forall f \in \mathcal{F} : R(f) \leq r(f, \mathbf{z}) + 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta, \quad (22)$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Теорема 5.7.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \forall f \in \mathcal{F} : R(f) \leq r(f, \mathbf{z}) + 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta, \quad (22)$$

$$P^n \left\{ \mathbf{z} : \forall f \in \mathcal{F} : R(f) \leq r(f, \mathbf{z}) + 2\mathcal{R}_n(\mathcal{F}, \mathbf{z}) + 3\sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (23)$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

◀ Заметим, что неравенство (22) является эквивалентной записью неравенства (17) из предыдущего утв. 5.11.

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

◀ Заметим, что неравенство (22) является эквивалентной записью неравенства (17) из предыдущего утв. 5.11.

Положим  $\delta' := \delta/2$ .



# Эмпирические и предсказывающие оценки

## Эмпирические оценки

◀ Заметим, что неравенство (22) является эквивалентной записью неравенства (17) из предыдущего утв. 5.11.

Положим  $\delta' := \delta/2$ .

Записывая для  $\delta'$  неравенство (22), получим

$$\mathbb{P}^n \left\{ \mathbf{z} : \forall f \in \mathcal{F} : R(f) \leq r(f, \mathbf{z}) + 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta'} \right)} \right\} \geq 1 - \delta'. \quad (24)$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Напоминание

**Утв. 5.10** Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ),  $\varepsilon > 0$  и  $n \in \mathbb{N}$ . Тогда

$$P^n \left\{ \mathbf{z} : \mathcal{R}_n(\mathcal{F}) - \mathcal{R}_n(\mathcal{F}, \mathbf{z}) \geq \varepsilon \right\} \leq \exp \left( \frac{-2n\varepsilon^2}{(b-a)^2} \right).$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

### Напоминание

**Утв. 5.10** Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [a, b]^Z$  ( $a < b$ ),  $\varepsilon > 0$  и  $n \in \mathbb{N}$ . Тогда

$$P^n \left\{ \mathbf{z} : \mathcal{R}_n(\mathcal{F}) - \mathcal{R}_n(\mathcal{F}, \mathbf{z}) \geq \varepsilon \right\} \leq \exp \left( \frac{-2n\varepsilon^2}{(b-a)^2} \right).$$

Из утв. 5.10 следует, что

$$P^n \left\{ \mathbf{z} : \mathcal{R}_n(\mathcal{F}) \leq \mathcal{R}_n(\mathcal{F}, \mathbf{z}) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta'} \right)} \right\} \geq 1 - \delta'. \quad (25)$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Объединяя вместе (24), (25) и применяя известную оценку для вероятности пересечения двух событий, получим

$$\mathbb{P}^n \left\{ \mathbf{z} : \forall f \in \mathcal{F} : R(f) \leq r(f, \mathbf{z}) + 2\mathcal{R}_n(\mathcal{F}, \mathbf{z}) + 3\sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta'} \right)} \right\} \geq 1 - 2\delta',$$

# Эмпирические и предсказывающие оценки

## Эмпирические оценки

Объединяя вместе (24), (25) и применяя известную оценку для вероятности пересечения двух событий, получим

$$\mathbb{P}^n \left\{ \mathbf{z} : \forall f \in \mathcal{F} : R(f) \leq r(f, \mathbf{z}) + 2\mathcal{R}_n(\mathcal{F}, \mathbf{z}) + 3\sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta'} \right)} \right\} \geq 1 - 2\delta',$$

а значит верно неравенство (23).



# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
  - Неравенство симметризации
  - Эмпирические оценки
  - **Предсказывающие оценки**
- 3 Оценка сверху для радемахеровской сложности

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Теорема 5.8.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Теорема 5.8.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда для любого метода минимизации эмпирического риска  $\mathbf{z} \mapsto f_{\mathbf{z}}$  ( $\mathbf{z} \in Z^n$ ) справедливо неравенство

$$P^n \left\{ \mathbf{z} : R(f_{\mathbf{z}}) \leq R(\mathcal{F}) + 4\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{2}{n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (26)$$



# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Теорема 5.8.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда для любого метода минимизации эмпирического риска  $\mathbf{z} \mapsto f_{\mathbf{z}}$  ( $\mathbf{z} \in Z^n$ ) справедливо неравенство

$$P^n \left\{ \mathbf{z} : R(f_{\mathbf{z}}) \leq R(\mathcal{F}) + 4\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{2}{n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (26)$$

◀ Для любых  $f \in \mathcal{F}$  и  $\mathbf{z} \in Z^n$  имеет место равенство

$$R(f_{\mathbf{z}}) = \left( R(f_{\mathbf{z}}) - r(f_{\mathbf{z}}, \mathbf{z}) \right) + \left( r(f_{\mathbf{z}}, \mathbf{z}) - r(f, \mathbf{z}) \right) + \left( r(f, \mathbf{z}) - R(f) \right) + R(f).$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

Откуда, учитывая произвольность выбора функции  $f \in \mathcal{F}$  и определение метода минимизации эмпирического риска, получим неравенство

$$R(f_{\mathbf{z}}) \leq \inf_{f \in \mathcal{F}} R(f) + \Delta_{\mathcal{F}}^+(\mathbf{z}) + \Delta_{\mathcal{F}}^-(\mathbf{z}). \quad (27)$$

Согласно утв. 5.11 имеют место оценки

$$\mathbf{P}^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \frac{\delta}{2}, \quad (28)$$

$$\mathbf{P}^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}^-(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \frac{\delta}{2}. \quad (29)$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

Объединяя вместе (27), (28) и (29), и применяя известное неравенство для вероятности пересечения событий, получим искомое неравенство (26).



# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Теорема 5.9.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Теорема 5.9.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (30)$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Теорема 5.9.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (30)$$

◀ Заметим, что для любого  $\varepsilon > 0$  имеет место равенство

$$P^n \{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) > \varepsilon \} = P^n \{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) > \varepsilon \} + P^n \{ \mathbf{z} : \Delta_{\mathcal{F}}^-(\mathbf{z}) > \varepsilon \}. \quad (31)$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Теорема 5.9.

Пусть  $P \in \mathcal{M}_+^1(Z)$ ,  $\mathcal{F} \subseteq [0, 1]^Z$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

Тогда

$$P^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) \leq 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (30)$$

◀ Заметим, что для любого  $\varepsilon > 0$  имеет место равенство

$$P^n \{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) > \varepsilon \} = P^n \{ \mathbf{z} : \Delta_{\mathcal{F}}^+(\mathbf{z}) > \varepsilon \} + P^n \{ \mathbf{z} : \Delta_{\mathcal{F}}^-(\mathbf{z}) > \varepsilon \}. \quad (31)$$

В равенстве (31) положим

$$\varepsilon := 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)}.$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

Применим неравенства (28) и (29) из доказательства предыдущей теоремы 5.8 к слагаемым, находящимся в правой части равенства (31).



# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

Применим неравенства (28) и (29) из доказательства предыдущей теоремы 5.8 к слагаемым, находящимся в правой части равенства (31).

Получим эквивалентное (30) неравенство

$$\mathbb{P}^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) > 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$



# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

Применим неравенства (28) и (29) из доказательства предыдущей теоремы 5.8 к слагаемым, находящимся в правой части равенства (31).

Получим эквивалентное (30) неравенство

$$\mathbb{P}^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) > 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$



По лемме 5.5 радемахеровская сложность классов функций  $\mathcal{H} \subseteq \{0, 1\}^X$  и  $\mathcal{F} \subseteq \{0, 1\}^Z$ , связанных соотношением  $\mathcal{H} \simeq_{l_01} \mathcal{F}$ , совпадает.

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

Применим неравенства (28) и (29) из доказательства предыдущей теоремы 5.8 к слагаемым, находящимся в правой части равенства (31).

Получим эквивалентное (30) неравенство

$$\mathbb{P}^n \left\{ \mathbf{z} : \Delta_{\mathcal{F}}(\mathbf{z}) > 2\mathcal{R}_n(\mathcal{F}) + \sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$



По лемме 5.5 радемахеровская сложность классов функций  $\mathcal{H} \subseteq \{0, 1\}^X$  и  $\mathcal{F} \subseteq \{0, 1\}^Z$ , связанных соотношением  $\mathcal{H} \simeq_{l_{01}} \mathcal{F}$ , совпадает.

Поэтому имеет место следующее утверждение, вытекающее из теорем 5.7 и 5.8.

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Следствие 5.2.

Пусть  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ .

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Следствие 5.2.

Пусть  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ . Тогда

$$\mathbb{P}^n \left\{ \mathbf{z} : \forall h \in \mathcal{H} : R(l_{01}; h) \leq r(l_{01}; h, \mathbf{z}) + \right. \\ \left. 2\mathcal{R}_n(\mathcal{H}, \mathbf{P}_X) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta, \quad (32)$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Следствие 5.2.

Пусть  $\mathcal{H} \subseteq \{0, 1\}^X$ ,  $\delta \in (0, 1)$  и  $n \in \mathbb{N}$ . Тогда

$$\mathbb{P}^n \left\{ \mathbf{z} : \forall h \in \mathcal{H} : R(l_{01}; h) \leq r(l_{01}; h, \mathbf{z}) + \right. \\ \left. 2\mathcal{R}_n(\mathcal{H}, \mathbf{P}_X) + \sqrt{\frac{1}{2n} \ln \left( \frac{1}{\delta} \right)} \right\} \geq 1 - \delta, \quad (32)$$

$$\mathbb{P}^n \left\{ \mathbf{z} : \forall h \in \mathcal{H} : R(l_{01}; h) \leq r(l_{01}; h, \mathbf{z}) + \right. \\ \left. 2\mathcal{R}_n(\mathcal{H}, \mathbf{z}|_X) + 3\sqrt{\frac{1}{2n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (33)$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Следствие 5.2 (продолжение).

Кроме того, для любого метода минимизации эмпирического риска  $\mathbf{z} \mapsto h_{\mathbf{z}}$  ( $\mathbf{z} \in Z^n$ ) справедливо неравенство

$$\mathbf{P}^n \left\{ \mathbf{z} : R(l_{01}; h_{\mathbf{z}}) \leq \inf_{h \in \mathcal{H}} R(l_{01}; h) + 4\mathcal{R}_n(\mathcal{H}, \mathbf{P}_X) + \sqrt{\frac{2}{n} \ln \left( \frac{2}{\delta} \right)} \right\} \geq 1 - \delta. \quad (34)$$

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Напоминание

**Лемма 5.6** Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$  и  $vc(\mathcal{F}) < \infty$ . Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \sqrt{\frac{2vc(\mathcal{F}) \ln(n+1)}{n}} \quad (\mathbf{z} \in Z^n).$$



# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Напоминание

**Лемма 5.6** Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$  и  $vc(\mathcal{F}) < \infty$ . Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \sqrt{\frac{2vc(\mathcal{F}) \ln(n+1)}{n}} \quad (\mathbf{z} \in Z^n).$$

Применяя лемму 5.6, получим следующий вариант неравенства Вапника-Червоненкиса.

### Следствие 5.3.

Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$  и  $vc(\mathcal{F}) < \infty$ .

# Эмпирические и предсказывающие оценки

## Предсказывающие оценки

### Напоминание

**Лемма 5.6** Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$  и  $\text{vc}(\mathcal{F}) < \infty$ . Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \sqrt{\frac{2\text{vc}(\mathcal{F}) \ln(n+1)}{n}} \quad (\mathbf{z} \in Z^n).$$

Применяя лемму 5.6, получим следующий вариант неравенства Вапника-Чевроненкиса.

### Следствие 5.3.

Пусть  $n \in \mathbb{N}$ ,  $\mathcal{F} \subseteq \{0, 1\}^Z$  и  $\text{vc}(\mathcal{F}) < \infty$ .

Тогда

$$\mathbf{E}_{\mathbf{z} \sim \mathcal{P}^n} [\Delta_{\mathcal{F}}^+(\mathbf{z})] \leq 2 \sqrt{\frac{2\text{vc}(\mathcal{F}) \ln(n+1)}{n}}. \quad (35)$$

# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности
  - $\epsilon$ -покрытия и  $\epsilon$ -упаковки
  - Теорема Дадли
  - Вероятностные оценки для бинарной классификации

# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности
  - $\epsilon$ -покрытия и  $\epsilon$ -упаковки
  - Теорема Дадли
  - Вероятностные оценки для бинарной классификации

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.9.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.9.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Множество  $A$  называется  **$\varepsilon$ -покрытием** множества  $B$ , если для любого  $b \in B$  найдется  $a \in A$  такое, что  $\rho(a, b) \leq \varepsilon$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.9.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Множество  $A$  называется  **$\varepsilon$ -покрытием** множества  $B$ , если для любого  $b \in B$  найдется  $a \in A$  такое, что  $\rho(a, b) \leq \varepsilon$ .

**Минимальным  $\varepsilon$ -покрытием** множества  $B$  называется  **$\varepsilon$ -покрытие** этого множества, обладающее наименьшей мощностью.

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.9.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Множество  $A$  называется  **$\varepsilon$ -покрытием** множества  $B$ , если для любого  $b \in B$  найдется  $a \in A$  такое, что  $\rho(a, b) \leq \varepsilon$ .

**Минимальным  $\varepsilon$ -покрытием** множества  $B$  называется  **$\varepsilon$ -покрытие** этого множества, обладающее наименьшей мощностью.

Мощность минимального  $\varepsilon$ -покрытия множества  $B$  называется **числом  $\varepsilon$ -покрытия**  $B$  и обозначается через  $N(\varepsilon, B, \rho)$ .



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.9.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Множество  $A$  называется  **$\varepsilon$ -покрытием** множества  $B$ , если для любого  $b \in B$  найдется  $a \in A$  такое, что  $\rho(a, b) \leq \varepsilon$ .

**Минимальным  $\varepsilon$ -покрытием** множества  $B$  называется  **$\varepsilon$ -покрытие** этого множества, обладающее наименьшей мощностью.

Мощность минимального  $\varepsilon$ -покрытия множества  $B$  называется **числом  $\varepsilon$ -покрытия**  $B$  и обозначается через  $N(\varepsilon, B, \rho)$ .

Будет использоваться также краткая запись  $N(\varepsilon, B)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.10.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.10.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Подмножество  $A \subseteq B$  называется  **$\varepsilon$ -упаковкой** множества  $B$ , если для любой  $a, b \in A$  выполняется условие  $\rho(a, b) > \varepsilon$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.10.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Подмножество  $A \subseteq B$  называется  **$\varepsilon$ -упаковкой** множества  $B$ , если для любой  $a, b \in A$  выполняется условие  $\rho(a, b) > \varepsilon$ .

**Максимальной  $\varepsilon$ -упаковкой** множества  $B$  называется  **$\varepsilon$ -упаковка** этого множества, обладающая наибольшей мощностью.

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.10.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Подмножество  $A \subseteq B$  называется  **$\varepsilon$ -упаковкой** множества  $B$ , если для любой  $a, b \in A$  выполняется условие  $\rho(a, b) > \varepsilon$ .

**Максимальной  $\varepsilon$ -упаковкой** множества  $B$  называется  **$\varepsilon$ -упаковка** этого множества, обладающая наибольшей мощностью.

Мощность максимальной  $\varepsilon$ -упаковки множества  $B$  называется **числом  $\varepsilon$ -упаковки**  $B$  и обозначается через  $M(\varepsilon, B, \rho)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.10.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Подмножество  $A \subseteq B$  называется  **$\varepsilon$ -упаковкой** множества  $B$ , если для любой  $a, b \in A$  выполняется условие  $\rho(a, b) > \varepsilon$ .

**Максимальной  $\varepsilon$ -упаковкой** множества  $B$  называется  **$\varepsilon$ -упаковка** этого множества, обладающая наибольшей мощностью.

Мощность максимальной  $\varepsilon$ -упаковки множества  $B$  называется **числом  $\varepsilon$ -упаковки**  $B$  и обозначается через  $M(\varepsilon, B, \rho)$ .

Будет использоваться также краткая запись  $M(\varepsilon, B)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Определение 5.10.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $A, B \subseteq M$  и  $\varepsilon > 0$ .

Подмножество  $A \subseteq B$  называется  **$\varepsilon$ -упаковкой** множества  $B$ , если для любой  $a, b \in A$  выполняется условие  $\rho(a, b) > \varepsilon$ .

**Максимальной  $\varepsilon$ -упаковкой** множества  $B$  называется  **$\varepsilon$ -упаковка** этого множества, обладающая наибольшей мощностью.

Мощность максимальной  $\varepsilon$ -упаковки множества  $B$  называется **числом  $\varepsilon$ -упаковки**  $B$  и обозначается через  $M(\varepsilon, B, \rho)$ .

Будет использоваться также краткая запись  $M(\varepsilon, B)$ .

Следующее утверждение устанавливает связь между введёнными понятиями.

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.7.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $B \subseteq M$  и  $\varepsilon > 0$ .



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.7.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $B \subseteq M$  и  $\varepsilon > 0$ .

Тогда

$$N(\varepsilon/2, B) \geq M(\varepsilon, B) \geq N(\varepsilon, B).$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.7.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $B \subseteq M$  и  $\varepsilon > 0$ .

Тогда

$$N(\varepsilon/2, B) \geq M(\varepsilon, B) \geq N(\varepsilon, B).$$

◀ Предположим, что  $A \subseteq B$  является максимальной  $\varepsilon$ -упаковкой множества  $B$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.7.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $B \subseteq M$  и  $\varepsilon > 0$ .

Тогда

$$N(\varepsilon/2, B) \geq M(\varepsilon, B) \geq N(\varepsilon, B).$$

◀ Предположим, что  $A \subseteq B$  является максимальной  $\varepsilon$ -упаковкой множества  $B$ .

Это означает, что для любого  $b \in B \setminus A$  множество  $A \cup \{b\}$  не обладает этим свойством.

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.7.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $B \subseteq M$  и  $\varepsilon > 0$ .

Тогда

$$N(\varepsilon/2, B) \geq M(\varepsilon, B) \geq N(\varepsilon, B).$$

◀ Предположим, что  $A \subseteq B$  является максимальной  $\varepsilon$ -упаковкой множества  $B$ .

Это означает, что для любого  $b \in B \setminus A$  множество  $A \cup \{b\}$  не обладает этим свойством.

Следовательно, найдётся  $a \in A$  такой, что  $\rho(a, b) \leq \varepsilon$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.7.

Пусть  $(M, \rho)$  – псевдометрическое пространство,  $B \subseteq M$  и  $\varepsilon > 0$ .

Тогда

$$N(\varepsilon/2, B) \geq M(\varepsilon, B) \geq N(\varepsilon, B).$$

◀ Предположим, что  $A \subseteq B$  является максимальной  $\varepsilon$ -упаковкой множества  $B$ .

Это означает, что для любого  $b \in B \setminus A$  множество  $A \cup \{b\}$  не обладает этим свойством.

Следовательно, найдётся  $a \in A$  такой, что  $\rho(a, b) \leq \varepsilon$ .

Таким образом,  $A$  одновременно является  $\varepsilon$ -покрытием множества  $B$ , а значит  $M(\varepsilon, B) = |A| \geq N(\varepsilon, B)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Предположим, что  $C \subseteq M$  является наименьшим  $\varepsilon/2$ -покрытием множества  $B$ , а  $A \subseteq B$  является  $\varepsilon$ -упаковкой этого множества.

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Предположим, что  $C \subseteq M$  является наименьшим  $\varepsilon/2$ -покрытием множества  $B$ , а  $A \subseteq B$  является  $\varepsilon$ -упаковкой этого множества.

Это означает, что для любого  $c \in C$  шар  $\{b \in B : \rho(c, b) \leq \varepsilon/2\}$  содержит не более одного элемента из  $A$ .

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Предположим, что  $C \subseteq M$  является наименьшим  $\varepsilon/2$ -покрытием множества  $B$ , а  $A \subseteq B$  является  $\varepsilon$ -упаковкой этого множества.

Это означает, что для любого  $c \in C$  шар  $\{b \in B : \rho(c, b) \leq \varepsilon/2\}$  содержит не более одного элемента из  $A$ .

Действительно, если предположить существование двух различных элементов  $a, a' \in A$ , одновременно принадлежащих такому шару, то из неравенства треугольника будет следовать  $\rho(a, a') \leq \rho(c, a') + \rho(c, a) \leq \varepsilon$ .



# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Предположим, что  $C \subseteq M$  является наименьшим  $\varepsilon/2$ -покрытием множества  $B$ , а  $A \subseteq B$  является  $\varepsilon$ -упаковкой этого множества.

Это означает, что для любого  $c \in C$  шар  $\{b \in B : \rho(c, b) \leq \varepsilon/2\}$  содержит не более одного элемента из  $A$ .

Действительно, если предположить существование двух различных элементов  $a, a' \in A$ , одновременно принадлежащих такому шару, то из неравенства треугольника будет следовать  $\rho(a, a') \leq \rho(c, a') + \rho(c, a) \leq \varepsilon$ .

А это нарушает свойство  $\varepsilon$ -упаковки, которым по предположению обладает множество  $A$ .

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Предположим, что  $C \subseteq M$  является наименьшим  $\varepsilon/2$ -покрытием множества  $B$ , а  $A \subseteq B$  является  $\varepsilon$ -упаковкой этого множества.

Это означает, что для любого  $c \in C$  шар  $\{b \in B : \rho(c, b) \leq \varepsilon/2\}$  содержит не более одного элемента из  $A$ .

Действительно, если предположить существование двух различных элементов  $a, a' \in A$ , одновременно принадлежащих такому шару, то из неравенства треугольника будет следовать  $\rho(a, a') \leq \rho(c, a') + \rho(c, a) \leq \varepsilon$ .

А это нарушает свойство  $\varepsilon$ -упаковки, которым по предположению обладает множество  $A$ .

Следовательно,  $|C| \geq |A|$ .

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Предположим, что  $C \subseteq M$  является наименьшим  $\varepsilon/2$ -покрытием множества  $B$ , а  $A \subseteq B$  является  $\varepsilon$ -упаковкой этого множества.

Это означает, что для любого  $c \in C$  шар  $\{b \in B : \rho(c, b) \leq \varepsilon/2\}$  содержит не более одного элемента из  $A$ .

Действительно, если предположить существование двух различных элементов  $a, a' \in A$ , одновременно принадлежащих такому шару, то из неравенства треугольника будет следовать  $\rho(a, a') \leq \rho(c, a') + \rho(c, a) \leq \varepsilon$ .

А это нарушает свойство  $\varepsilon$ -упаковки, которым по предположению обладает множество  $A$ .

Следовательно,  $|C| \geq |A|$ .

В виду произвольности выбора  $\varepsilon$ -упаковки  $A$  получим неравенство  $N(\varepsilon/2, B) \geq M(\varepsilon, B)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

В дальнейшем нас будет интересовать специальный случай  
псевдометрических пространств,

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

В дальнейшем нас будет интересовать специальный случай псевдометрических пространств,

у которых в качестве множества элементов выступают классы функций  $\mathcal{F} \subseteq \mathbb{R}^Z$ ,

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

В дальнейшем нас будет интересовать специальный случай псевдометрических пространств,

у которых в качестве множества элементов выступают классы функций  $\mathcal{F} \subseteq \mathbb{R}^Z$ ,

а псевдометрики порождаются полунормами вида

$$\|f\|_{p,\mathbf{z}} := \left( \frac{1}{n} \sum_{i=1}^n |f(z_i)|^p \right)^{1/p} \quad (f \in \mathcal{F}),$$

где  $p \in [1, \infty)$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Утверждение 5.12.

Пусть  $a > 0$ ,  $b \geq 2e$  и  $a \leq b \ln(a)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Утверждение 5.12.

Пусть  $a > 0$ ,  $b \geq 2e$  и  $a \leq b \ln(a)$ .

Тогда

$$a \leq \frac{e}{e-1} b \ln(b). \quad (36)$$



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Утверждение 5.12.

Пусть  $a > 0$ ,  $b \geq 2e$  и  $a \leq b \ln(a)$ .

Тогда

$$a \leq \frac{e}{e-1} b \ln(b). \quad (36)$$

◀ Обозначим  $\gamma := \frac{e}{e-1}$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Утверждение 5.12.

Пусть  $a > 0$ ,  $b \geq 2e$  и  $a \leq b \ln(a)$ .

Тогда

$$a \leq \frac{e}{e-1} b \ln(b). \quad (36)$$

◀ Обозначим  $\gamma := \frac{e}{e-1}$ .

Для доказательства этого утверждения достаточно показать, что  $a \geq b \ln(a)$  при  $a \geq \gamma b \ln(b)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Утверждение 5.12.

Пусть  $a > 0$ ,  $b \geq 2e$  и  $a \leq b \ln(a)$ .

Тогда

$$a \leq \frac{e}{e-1} b \ln(b). \quad (36)$$

◀ Обозначим  $\gamma := \frac{e}{e-1}$ .

Для доказательства этого утверждения достаточно показать, что  $a \geq b \ln(a)$  при  $a \geq \gamma b \ln(b)$ .

С этой целью исследуем поведение функции  $\varphi(a) := a - b \ln(a)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Утверждение 5.12.

Пусть  $a > 0$ ,  $b \geq 2e$  и  $a \leq b \ln(a)$ .

Тогда

$$a \leq \frac{e}{e-1} b \ln(b). \quad (36)$$

◀ Обозначим  $\gamma := \frac{e}{e-1}$ .

Для доказательства этого утверждения достаточно показать, что  $a \geq b \ln(a)$  при  $a \geq \gamma b \ln(b)$ .

С этой целью исследуем поведение функции  $\varphi(a) := a - b \ln(a)$ .

Производная этой функции  $\varphi'(a) = 1 - b/a > 0$  при  $a \in (b, +\infty)$ ,

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Утверждение 5.12.

Пусть  $a > 0$ ,  $b \geq 2e$  и  $a \leq b \ln(a)$ .

Тогда

$$a \leq \frac{e}{e-1} b \ln(b). \quad (36)$$

◀ Обозначим  $\gamma := \frac{e}{e-1}$ .

Для доказательства этого утверждения достаточно показать, что  $a \geq b \ln(a)$  при  $a \geq \gamma b \ln(b)$ .

С этой целью исследуем поведение функции  $\varphi(a) := a - b \ln(a)$ .

Производная этой функции  $\varphi'(a) = 1 - b/a > 0$  при  $a \in (b, +\infty)$ , а это означает, что сама функция  $\varphi(a)$  возрастает на этом промежутке.

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Заметим, что  $\gamma b \ln(b) \in (b, +\infty)$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Заметим, что  $\gamma b \ln(b) \in (b, +\infty)$ .

## Замечание

$e/(e-1) = 1.5819767068693264243850020051090115585468693010753961362667870596\dots$

Действительно,  $\gamma b \ln(b) > b \ln(2e) > b$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Заметим, что  $\gamma b \ln(b) \in (b, +\infty)$ .

## Замечание

$e/(e-1) = 1.5819767068693264243850020051090115585468693010753961362667870596\dots$

Действительно,  $\gamma b \ln(b) > b \ln(2e) > b$ .

Поэтому нам достаточно будет показать, что  $\varphi(\gamma b \ln(b)) \geq 0$ .



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Заметим, что  $\gamma b \ln(b) \in (b, +\infty)$ .

## Замечание

$e/(e-1) = 1.5819767068693264243850020051090115585468693010753961362667870596\dots$

Действительно,  $\gamma b \ln(b) > b \ln(2e) > b$ .

Поэтому нам достаточно будет показать, что  $\varphi(\gamma b \ln(b)) \geq 0$ .

Проверим это

$$\begin{aligned}\varphi(\gamma b \ln(b)) &= \gamma b \ln(b) - b \ln(\gamma b \ln(b)) = \\ &= b[(\gamma - 1) \ln(b) - \ln(\gamma \ln(b))] = \\ &= b[e^{-1}c - \ln(c)],\end{aligned}$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Заметим, что  $\gamma b \ln(b) \in (b, +\infty)$ .

## Замечание

$e/(e-1) = 1.5819767068693264243850020051090115585468693010753961362667870596\dots$

Действительно,  $\gamma b \ln(b) > b \ln(2e) > b$ .

Поэтому нам достаточно будет показать, что  $\varphi(\gamma b \ln(b)) \geq 0$ .

Проверим это

$$\begin{aligned}\varphi(\gamma b \ln(b)) &= \gamma b \ln(b) - b \ln(\gamma b \ln(b)) = \\ &= b[(\gamma - 1) \ln(b) - \ln(\gamma \ln(b))] = \\ &= b[e^{-1}c - \ln(c)],\end{aligned}$$

где  $c := \gamma \ln(b)$  и  $c \geq \gamma \ln(2e) > 0$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Покажем, что функции  $\psi(c) := e^{-1}c - \ln(c) \geq 0$  при  $c > 0$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Покажем, что функции  $\psi(c) := e^{-1}c - \ln(c) \geq 0$  при  $c > 0$ .

Производная этой функции  $\psi(c) = e^{-1} - 1/c$  обращается в нуль при  $c = e$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Покажем, что функции  $\psi(c) := e^{-1}c - \ln(c) \geq 0$  при  $c > 0$ .

Производная этой функции  $\psi(c) = e^{-1} - 1/c$  обращается в нуль при  $c = e$ .

Это точка минимума функции  $\psi(c)$ , в которой  $\psi(e) = 0$ .



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.8.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.8.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{1,\mathbf{z}}) \leq \left( \frac{9}{\varepsilon} \ln \left( \frac{2e}{\varepsilon} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (37)$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.8.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{1,\mathbf{z}}) \leq \left( \frac{9}{\varepsilon} \ln \left( \frac{2e}{\varepsilon} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (37)$$

◀ Для доказательства данного утверждения воспользуемся так называемым вероятностным методом.



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.8.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{1,\mathbf{z}}) \leq \left( \frac{9}{\varepsilon} \ln \left( \frac{2e}{\varepsilon} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (37)$$

◀ Для доказательства данного утверждения воспользуемся так называемым вероятностным методом.

Вначале выполним вспомогательные построения.

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.8.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{1,\mathbf{z}}) \leq \left( \frac{9}{\varepsilon} \ln \left( \frac{2e}{\varepsilon} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (37)$$

◀ Для доказательства данного утверждения воспользуемся так называемым вероятностным методом.

Вначале выполним вспомогательные построения.

Обозначим множество индексов  $\Omega_n := \{1, \dots, n\}$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

## Лемма 5.8.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{1, \mathbf{z}}) \leq \left( \frac{9}{\varepsilon} \ln \left( \frac{2e}{\varepsilon} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (37)$$

◀ Для доказательства данного утверждения воспользуемся так называемым вероятностным методом.

Вначале выполним вспомогательные построения.

Обозначим множество индексов  $\Omega_n := \{1, \dots, n\}$

и определим вероятностное пространство  $(\Omega_n, 2^{\Omega_n}, \mathbf{U})$  с равномерно распределённой на  $\Omega_n$  вероятностной мерой  $\mathbf{U}$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Зафиксируем  $m \in \mathbb{N}$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Зафиксируем  $m \in \mathbb{N}$ .

На вероятностном пространстве  $(\Omega_n^m, \otimes^m 2^{\Omega_n}, U^m)$  функции координатных проекций  $I_1, \dots, I_m$  являются независимыми случайными величинами с распределением  $U$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Зафиксируем  $m \in \mathbb{N}$ .

На вероятностном пространстве  $(\Omega_n^m, \otimes^m 2^{\Omega_n}, U^m)$  функции координатных проекций  $I_1, \dots, I_m$  являются независимыми случайными величинами с распределением  $U$ .

Пусть  $\{f_1, \dots, f_M\} \subseteq \mathcal{F}$ , где  $M = M(\varepsilon, \mathcal{F}, \|\cdot\|_{1,z})$ , – максимальная  $\varepsilon$ -упаковка множества  $\mathcal{F}$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Зафиксируем  $m \in \mathbb{N}$ .

На вероятностном пространстве  $(\Omega_n^m, \otimes^m 2^{\Omega_n}, U^m)$  функции координатных проекций  $I_1, \dots, I_m$  являются независимыми случайными величинами с распределением  $U$ .

Пусть  $\{f_1, \dots, f_M\} \subseteq \mathcal{F}$ , где  $M = M(\varepsilon, \mathcal{F}, \|\cdot\|_{1,z})$ , – максимальная  $\varepsilon$ -упаковка множества  $\mathcal{F}$ .

Для любой пары индексов  $k, l \in \Omega_n$ ,  $k \neq l$  выполняется неравенство

$$U\{i : f_k(z_i) \neq f_l(z_i)\} = \|f_k - f_l\|_{1,z} > \varepsilon,$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

но тогда

$$\begin{aligned} U^m \{f_k(z_{I_j}) = f_l(z_{I_j}) \ (j = 1, \dots, m)\} &= \prod_{j=1}^m U^m \{f_k(z_{I_j}) = f_l(z_{I_j})\} \\ &\leq (1 - \varepsilon)^m \leq \left| \begin{array}{l} 1 + u \leq e^u \\ (u \in \mathbb{R}) \end{array} \right| \\ &\leq e^{-m\varepsilon}. \end{aligned}$$



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

но тогда

$$\begin{aligned} U^m \{f_k(z_{I_j}) = f_l(z_{I_j}) \ (j = 1, \dots, m)\} &= \prod_{j=1}^m U^m \{f_k(z_{I_j}) = f_l(z_{I_j})\} \\ &\leq (1 - \varepsilon)^m \leq \left| \begin{array}{l} 1 + u \leq e^u \\ (u \in \mathbb{R}) \end{array} \right| \\ &\leq e^{-m\varepsilon}. \end{aligned}$$

Используя неравенство для вероятности объединения событий, получим

$$U^m \{ \exists k, l : k \neq l \text{ и } f_k(z_{I_j}) = f_l(z_{I_j}) \ (j = 1, \dots, m) \} < M^2 e^{-m\varepsilon}. \quad (38)$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

но тогда

$$\begin{aligned} U^m \{f_k(z_{I_j}) = f_l(z_{I_j}) \ (j = 1, \dots, m)\} &= \prod_{j=1}^m U^m \{f_k(z_{I_j}) = f_l(z_{I_j})\} \\ &\leq (1 - \varepsilon)^m \leq \left| \begin{array}{l} 1 + u \leq e^u \\ (u \in \mathbb{R}) \end{array} \right| \\ &\leq e^{-m\varepsilon}. \end{aligned}$$

Используя неравенство для вероятности объединения событий, получим

$$U^m \{ \exists k, l : k \neq l \text{ и } f_k(z_{I_j}) = f_l(z_{I_j}) \ (j = 1, \dots, m) \} < M^2 e^{-m\varepsilon}. \quad (38)$$

Выберем  $m := \frac{2}{\varepsilon} \ln(M)$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

но тогда

$$\begin{aligned} \mathbb{U}^m \{f_k(z_{l_j}) = f_l(z_{l_j}) \ (j = 1, \dots, m)\} &= \prod_{j=1}^m \mathbb{U}^m \{f_k(z_{l_j}) = f_l(z_{l_j})\} \\ &\leq (1 - \varepsilon)^m \leq \left| \begin{array}{l} 1 + u \leq e^u \\ (u \in \mathbb{R}) \end{array} \right| \\ &\leq e^{-m\varepsilon}. \end{aligned}$$

Используя неравенство для вероятности объединения событий, получим

$$\mathbb{U}^m \{ \exists k, l : k \neq l \text{ и } f_k(z_{l_j}) = f_l(z_{l_j}) \ (j = 1, \dots, m) \} < M^2 e^{-m\varepsilon}. \quad (38)$$

Выберем  $m := \frac{2}{\varepsilon} \ln(M)$

и заметим, что в этом случае вероятность события в левой части неравенства (38) меньше единицы.

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Это, в свою очередь, означает существование последовательности индексов  $(i_1, \dots, i_m) \in \Omega_n^m$  такой, что на множестве  $\{z_{i_1}, \dots, z_{i_m}\}$  функции  $f_1, \dots, f_M$  попарно отличаются друг от друга.

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Это, в свою очередь, означает существование последовательности индексов  $(i_1, \dots, i_m) \in \Omega_n^m$  такой, что на множестве  $\{z_{i_1}, \dots, z_{i_m}\}$  функции  $f_1, \dots, f_M$  попарно отличаются друг от друга.

Но тогда

$$M \leq \Gamma_{\mathcal{F}}(m). \quad (39)$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Это, в свою очередь, означает существование последовательности индексов  $(i_1, \dots, i_m) \in \Omega_n^m$  такой, что на множестве  $\{z_{i_1}, \dots, z_{i_m}\}$  функции  $f_1, \dots, f_M$  попарно отличаются друг от друга.

Но тогда

$$M \leq \Gamma_{\mathcal{F}}(m). \quad (39)$$

Далее, возможны два случая.

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Это, в свою очередь, означает существование последовательности индексов  $(i_1, \dots, i_m) \in \Omega_n^m$  такой, что на множестве  $\{z_{i_1}, \dots, z_{i_m}\}$  функции  $f_1, \dots, f_M$  попарно отличаются друг от друга.

Но тогда

$$M \leq \Gamma_{\mathcal{F}}(m). \quad (39)$$

Далее, возможны два случая.

Для краткости обозначим  $d := \text{vc}(\mathcal{F})$ .

# Оценка сверху для радемахеровской сложности

## $\varepsilon$ -покрытия и $\varepsilon$ -упаковки

Это, в свою очередь, означает существование последовательности индексов  $(i_1, \dots, i_m) \in \Omega_n^m$  такой, что на множестве  $\{z_{i_1}, \dots, z_{i_m}\}$  функции  $f_1, \dots, f_M$  попарно отличаются друг от друга.

Но тогда

$$M \leq \Gamma_{\mathcal{F}}(m). \quad (39)$$

Далее, возможны два случая.

Для краткости обозначим  $d := \text{vc}(\mathcal{F})$ .

Если  $m \leq d$ , то  $\Gamma_{\mathcal{F}}(m) = 2^d$  и неравенство (39) принимает вид  $M \leq 2^d$ .



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Это, в свою очередь, означает существование последовательности индексов  $(i_1, \dots, i_m) \in \Omega_n^m$  такой, что на множестве  $\{z_{i_1}, \dots, z_{i_m}\}$  функции  $f_1, \dots, f_M$  попарно отличаются друг от друга.

Но тогда

$$M \leq \Gamma_{\mathcal{F}}(m). \quad (39)$$

Далее, возможны два случая.

Для краткости обозначим  $d := \text{vc}(\mathcal{F})$ .

Если  $m \leq d$ , то  $\Gamma_{\mathcal{F}}(m) = 2^d$  и неравенство (39) принимает вид  $M \leq 2^d$ .

Откуда сразу вытекает искомое неравенство (37).

## Замечание

$9 \ln(2e) = 15.238324625039507784755089093123589112679501209242297287086120085\dots$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Если  $m > d$ , то с учётом леммы Сауэра-Шелаха неравенство (39) может быть преобразовано к виду

$$M_d^1 \leq \frac{em}{d} = \frac{2e}{\varepsilon} \ln(M_d^1).$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Если  $m > d$ , то с учётом леммы Сауэра-Шелаха неравенство (39) может быть преобразовано к виду

$$M_d^1 \leq \frac{em}{d} = \frac{2e}{\varepsilon} \ln(M_d^1).$$

Но тогда выполняется неравенство (36) из предыдущего утв. 5.12, в котором  $a = M_d^1$  и  $b = \frac{em}{d}$ .

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Если  $m > d$ , то с учётом леммы Сауэра-Шелаха неравенство (39) может быть преобразовано к виду

$$M_d^{\frac{1}{d}} \leq \frac{em}{d} = \frac{2e}{\varepsilon} \ln(M_d^{\frac{1}{d}}).$$

Но тогда выполняется неравенство (36) из предыдущего утв. 5.12, в котором  $a = M_d^{\frac{1}{d}}$  и  $b = \frac{em}{d}$ .

Это значит

$$M \leq \left( \frac{e}{e-1} \frac{2e}{\varepsilon} \ln \left( \frac{2e}{\varepsilon} \right) \right)^d.$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Если  $m > d$ , то с учётом леммы Сауэра-Шелаха неравенство (39) может быть преобразовано к виду

$$M_d^1 \leq \frac{em}{d} = \frac{2e}{\varepsilon} \ln(M_d^1).$$

Но тогда выполняется неравенство (36) из предыдущего утв. 5.12, в котором  $a = M_d^1$  и  $b = \frac{em}{d}$ .

Это значит

$$M \leq \left( \frac{e}{e-1} \frac{2e}{\varepsilon} \ln \left( \frac{2e}{\varepsilon} \right) \right)^d.$$

Учитывая оценку  $\frac{2e^2}{e-1} < 9$ , получим требуемое неравенство (37).



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Обобщим только что полученный результат.

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Обобщим только что полученный результат.

## Лемма 5.9.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $p \in [1, \infty)$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Обобщим только что полученный результат.

## Лемма 5.9.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $p \in [1, \infty)$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{p, \mathbf{z}}) \leq \left( \frac{9}{\varepsilon^p} \ln \left( \frac{2e}{\varepsilon^p} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (40)$$



# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Обобщим только что полученный результат.

## Лемма 5.9.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $p \in [1, \infty)$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{p, \mathbf{z}}) \leq \left( \frac{9}{\varepsilon^p} \ln \left( \frac{2e}{\varepsilon^p} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (40)$$

◀ Заметим, что в рассматриваемом случае выполняются равенства

$$\|f_1 - f_2\|_{p, \mathbf{z}}^p = \|f_1 - f_2\|_{1, \mathbf{z}} \quad (f_1, f_2 \in \mathcal{F}),$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Обобщим только что полученный результат.

## Лемма 5.9.

Пусть  $\mathcal{F} \subseteq \{0, 1\}^Z$ ,  $\varepsilon \in (0, 1]$ ,  $p \in [1, \infty)$ ,  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Тогда

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{p, \mathbf{z}}) \leq \left( \frac{9}{\varepsilon^p} \ln \left( \frac{2e}{\varepsilon^p} \right) \right)^{\text{vc}(\mathcal{F})}. \quad (40)$$

◀ Заметим, что в рассматриваемом случае выполняются равенства

$$\|f_1 - f_2\|_{p, \mathbf{z}}^p = \|f_1 - f_2\|_{1, \mathbf{z}} \quad (f_1, f_2 \in \mathcal{F}),$$

а значит

$$M(\varepsilon, \mathcal{F}, \|\cdot\|_{p, \mathbf{z}}) = M(\varepsilon^p, \mathcal{F}, \|\cdot\|_{1, \mathbf{z}}). \quad (41)$$

# Оценка сверху для радемахеровской сложности

$\varepsilon$ -покрытия и  $\varepsilon$ -упаковки

Объединяя (37) из предыдущей леммы 5.8 и (41), получим искомое неравенство (40).



# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности
  - $\epsilon$ -покрытия и  $\epsilon$ -упаковки
  - Теорема Дадли
  - Вероятностные оценки для бинарной классификации

# Оценка сверху для радемахеровской сложности

Теорема Дадли

Теорема 5.10 (Дадли).

Пусть  $\mathcal{F} \subseteq \mathbb{R}^Z$  и  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

# Оценка сверху для радемахеровской сложности

Теорема Дадли

Теорема 5.10 (Дадли).

Пусть  $\mathcal{F} \subseteq \mathbb{R}^Z$  и  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Предположим, что

$$\gamma_0 := \sup_{f \in \mathcal{F}} \|f\|_{2, \mathbf{z}} < \infty.$$

# Оценка сверху для радемахеровской сложности

Теорема Дадли

## Теорема 5.10 (Дадли).

Пусть  $\mathcal{F} \subseteq \mathbb{R}^Z$  и  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Предположим, что

$$\gamma_0 := \sup_{f \in \mathcal{F}} \|f\|_{2, \mathbf{z}} < \infty.$$

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \frac{12}{\sqrt{n}} \int_0^{\gamma_0} \sqrt{\ln N(\varepsilon, \mathcal{F}, \|\cdot\|_{2, \mathbf{z}})} d\varepsilon. \quad (42)$$

# Оценка сверху для радемахеровской сложности

Теорема Дадли

## Теорема 5.10 (Дадли).

Пусть  $\mathcal{F} \subseteq \mathbb{R}^Z$  и  $\mathbf{z} \in Z^n$  ( $n \in \mathbb{N}$ ).

Предположим, что

$$\gamma_0 := \sup_{f \in \mathcal{F}} \|f\|_{2, \mathbf{z}} < \infty.$$

Тогда

$$\mathcal{R}_n(\mathcal{F}, \mathbf{z}) \leq \frac{12}{\sqrt{n}} \int_0^{\gamma_0} \sqrt{\ln N(\varepsilon, \mathcal{F}, \|\cdot\|_{2, \mathbf{z}})} d\varepsilon. \quad (42)$$





# Оценка сверху для радемахеровской сложности

Теорема Дадли

# Оценка сверху для радемахеровской сложности

Теорема Дадли

# Оценка сверху для радемахеровской сложности

Теорема Дадли

# Оценка сверху для радемахеровской сложности

Теорема Дадли

# Содержание

- 1 Радемахеровская сложность
- 2 Эмпирические и предсказывающие оценки
- 3 Оценка сверху для радемахеровской сложности
  - $\epsilon$ -покрытия и  $\epsilon$ -упаковки
  - Теорема Дадли
  - Вероятностные оценки для бинарной классификации

# Оценка сверху для радемахеровской сложности

Вероятностные оценки для бинарной классификации

# Оценка сверху для радемахеровской сложности

Вероятностные оценки для бинарной классификации

# Оценка сверху для радемахеровской сложности

Вероятностные оценки для бинарной классификации



# Оценка сверху для радемахеровской сложности

Вероятностные оценки для бинарной классификации

# Оценка сверху для радемахеровской сложности

Вероятностные оценки для бинарной классификации