ЛЕКЦИЯ 22

ТЕОРЕМА ГЛИВЕНКО-КАНТЕЛЛИ. МАТРИЦЫ АДАМАРА

1. Теорема Гливенко – Кантелли

В этой лекции будут рассмотрены приложения к статистике. Сначала рассмотрим классическое утверждение из теории вероятности:

Утверждение 19 (Закон больших чисел для схемы испытаний Бернулли) Пусть задана $\xi_1, \dots, \xi_n, \dots$ — последовательность случайных величин, где каждая величина:

$$\xi_i = \begin{cases} 1 \text{ с вероятностью } p, \\ 0 \text{ с вероятностью } q = 1 - p. \end{cases}$$

Пусть также эти величины — независимы в совокупности. Тогда:

$$\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{P} M\xi_1 = p, \quad n \to \infty.$$

Утверждение 20 (Усиленный закон больших чисел) Пусть выполняются условия закона больших чисел для схемы испытаний Бернулли. Тогда:

$$\frac{\xi_1 + \dots + \xi_n}{n} \to P \text{ почти наверное}, \quad n \to \infty.$$

Замечание На усиленном законе больших чисел основывается, в том числе, метод Монте – Карло. $*$

Определение 40: ***Эмпирической функцией распределения*** $\hat{F}_n$ *называется следующее:*

$$\hat{F}_n(\xi_1, \dots, \xi_n; x) = \frac{1}{n} \sum_{i=1}^{n} I_{\{\xi_i \leqslant x\}}.$$

Обозначим:

$$I_{\{\xi_i \leqslant x\}} = \eta_i,$$

тогда:

$$\widehat{F}_n(\xi_1, \dots, \xi_n; x) = \frac{\eta_1 + \dots + \eta_n}{n}.$$

Заметим, что:

$$\eta_i = \begin{cases} 1 \text{ с вероятностью } p = F(x), \\ 0 \text{ с вероятностью } q = 1 - p. \end{cases}$$

Из-за независимости исходных величин ξ_i также независимы величины η_i.

Теорема 67 (Гливенко – Кантелли)

$$\sup_{x \in \mathbb{R}} \left| \widehat{F}_n(\xi_1, \dots, \xi_n; x) - F(x) \right| \to 0 \text{ почти наверное}, \quad n \to \infty.$$

Замечание Теорема Гливенко – Кантелли фактически является утверждением о равномерной по всем $x \in \mathbb{R}$ сходимости в усиленном законе больших чисел для специальной схемы испытаний Бернулли, заданной эмпирической формулой распределения. $*$

Поставим теперь общую задачу о равномерной сходимости в усиленном законе больших чисел для схемы Бернулли.

Пусть есть семейство последовательностей событий на некотором вероятностном пространстве:

$$A_1^x, A_2^x, \dots, A_n^x, \dots, \quad x \in \mathcal{X}.$$

Например, в задаче Гливенко – Кантелли семейством последовательностей является:

$$A_i^x = \{\xi_i \leqslant x\}, \quad \mathcal{X} = \mathbb{R}.$$

В общем случае предположим также следующее:

$$\forall x \in \mathcal{X} \text{ все } A_i^x \text{ — независимы в совокупности.}$$

Далее предположим, что:

$$\forall x \in \mathcal{X}, \ \forall i \quad P(A_i^x) = p^x.$$

Тогда возникает вопрос — при каких условиях можно гарантировать, что усиленный закон больших чисел выполнен равномерно по всем $x \in \mathbb{R}$? Запишем поэтому:

$$\sup_{x \in \mathbb{R}} \left| \frac{I_{A_1^x} + \dots + I_{A_n^x}}{n} - p^x \right| \to 0, \quad n \to \infty.$$

Теорема 68 Равномерная сходимость в усиленном законе больших чисел имеет место тогда и только тогда, когда конечна размерность Вапника – Червоненкиса ранжированного пространства, в котором:

$$S = \Omega, \quad R \text{ — это все } A_i^x.$$

2. Матрицы Адамара

Рассмотрим классический объект в комбинаторике, связанный со многими уже изученными объектами:

Определение 41: *Матрица A размера $n \times n$ называется* ***матрицей Адамара****, если она состоит только из «-1» и «1», причем любые две ее строки должны быть ортогональны в евклидовом смысле.* ♣

Заметим, что при $n = 1$ матрица Адамара существует, и даже две — одна состоит только эз элемента «-1», а другая — из элемента «1».

При $n = 2$ тоже очевидно существование матрицы Адамара:

$$\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$

При $n = 3$ не существует матрицы Адамара, более того — матрицы Адамара не существуют при любых нечетных n, больших единицы.

Это объясняется тем, что если строк в матрице хотя бы две, и у каждой из них нечетная длина, то их скалярное произведение — нечетное число, то есть точно не ноль.

Пусть есть какая-то матрица Адамара при $n > 1$. Если умножить любые ее строки на «-1», то снова получится матрица Адамара.

Более того, если умножить столбцы матрица Адамара при $n > 1$ на «-1» — тоже получится матрица Адамара.

Задача Доказать, что в определении 41 попарную ортогональность строк можно заменить попарной ортогональностью столбцов. $*$

В результате можно считать, что матрица Адамара имеет нормальную форму, то есть:

$$\begin{pmatrix} 1 & 1 & & \ldots\ldots & & & 1 \\ 1 & -1 & \ldots & -1 & 1 & \ldots & 1 \\ 1 & & & & & & \\ \vdots & & & \ddots & & & \\ 1 & & & & & & \end{pmatrix}$$

В этой матрице первая строчка целиком состоит из «1», во второй строчке сначала стоит «1», а затем идут половина «-1» и половина «1», и так далее.

Отсюда и вытекает очевидное ограничение — матрицы Адамара не существуют при нечетных n, больших единицы.

Есть еще одно очевидное ограничение на существование матрицы Адамара:

Утверждение 21

$$n \equiv 0(4), \quad n \geqslant 3.$$

Тогда для ортогональности строк матрица должна иметь следующий вид:

$$\begin{pmatrix} 1 & & \ldots\ldots & & 1 & -1 & & \ldots\ldots & & -1 \\ 1 & \ldots & 1 & -1 & \ldots & -1 & 1 & \ldots & 1 & -1 & \ldots & -1 \\ & & & & \ldots\ldots\ldots\ldots\ldots\ldots & & & & & & \end{pmatrix}$$

Задача Доказать утверждение 21. $*$

Существует следующая **гипотеза** — если $n > 2$ и $n \equiv 0(4)$, то существует матрица Адамара.

Вплоть до $n = 668$ гипотеза точно верна. Более того, до сих пор неизвестен такой n, при котором данная гипотеза была бы неверна.

Теорема 69

$$\exists f: \ f(n) = \overline{\overline{o}}(n): \ \forall n \text{ на отрезке от } n \text{ до } n + f(n)$$

имеется порядок матрицы Адамара. $*$

То есть, порядки матрицы Адамара так же плотны в натуральном ряде, как и простые числа.

Применим **наивный подход** к построению матрицы Адамара. Пусть:

$$n = 2^k.$$

Тогда:

$$\begin{pmatrix} 1 & & & \multicolumn{7}{c}{\dots\dots\dots\dots\dots\dots} & & 1 \\ 1 & & \multicolumn{3}{c}{\dots\dots} & 1 & -1 & \multicolumn{3}{c}{\dots\dots} & -1 \\ 1 & \dots & 1 & -1 & \dots & -1 & 1 & \dots & 1 & -1 & \dots & -1 \\ & & \multicolumn{8}{c}{\dots\dots\dots\dots\dots\dots\dots} & & \end{pmatrix}$$

В верхней строке записано n единиц, во второй строке — $\frac{n}{2}$ «1» и $\frac{n}{2}$ «-1» в указанном порядке. Видно, что вторая строка ортогональна первой.

В третью строку записаны в следующем порядке: $\frac{n}{4}$ «1», $\frac{n}{4}$ «-1», $\frac{n}{4}$ «1», $\frac{n}{4}$ «-1», и так далее, так как n делится на любую степень двойки.

В результате строчек всего получится:

$$k = \log_2 n,$$

а не n, которое нужно было получить.

То есть, наивный подход не работает применительно к данной проблеме.

Известен подход для тех же $n = 2^k$, который работает. Продемонстрируем этот подход на матрице размера 4×4:

$$\begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Видно, что получена матрица Адамара.

Этот принцип построения верен всегда. Заметим, что при построении использовалось понятие кронекеровского произведения матриц:

Определение 42: Кронекеровским произведением матриц $A * B$ *называется следующая матрица:*

$$A * B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \dots & a_{nn}B \end{pmatrix}$$

Утверждение 22 Если A и B являются матрицами Адамара, то их кронекеровское произведение также является матрицей Адамара. $*$

Задача Доказать утверждение 22. $*$

Таким образом, получена бесконечная серия примеров матриц Адамара.

Матрица Адамара является важным объектом в теории кодирования. Тем не менее, сейчас матрицы Адамара будут рассмотрены в связи с изученным понятием *дистанционного графа.*

Дистанционные графы ранее возникали при изучении задачи о хроматическом числе, в проблеме Борсука и даже в исследовании чисел Рамсея.

Рассматривались графы вида $G(n, r, s)$. В частности, значительную роль в комбинаторной геометрии играет граф вида:

$$G\left(n, \frac{n}{2}, \frac{n}{4}\right), \quad n \equiv 0(4).$$

Из существования графа вида $G\left(n, \frac{n}{2}, \frac{n}{4}\right)$ следовала следующая оценка:

$$\chi(\mathbb{R}^n) \geqslant \chi\left(G\left(n, \frac{n}{2}, \frac{n}{4}\right)\right) \geqslant \frac{C_n^{\frac{n}{2}}}{\alpha\left(G\left(n, \frac{n}{2}, \frac{n}{4}\right)\right)}.$$

В частности, было рассмотрено число независимости:

$$\alpha\left(G\left(n, \frac{n}{2}, \frac{n}{4}\right)\right).$$

Оказывается, что:

Утверждение 23 Максимальная клика в $G\left(n, \frac{n}{2}, \frac{n}{4}\right)$ является матрицей Адамара. $*$

Док-во: Возьмем матрицу Адамара, и заменим все «-1» на «0». Тогда условие попарной ортогональности изменится на то, что попарные скалярные произведения будут равны $\frac{n}{4}$.

Отсюда следует, что вершины попарно соединены ребрами, а значит, образуют клику.

Максимальность клики доказывается тривиальными соображениями, основанными на линейно-алгебраическом методе. ■

Заметим, что полученная ситуация не является той же самой, которая использовалась ранее при оценке диагонального числа Рамсея.

На следующей лекции будет рассмотрено приложение теории о матрицах Адамара.