07/02/2020 - High-dimensional probability. An Introduction with applications in Data Science (Vershynin R.)

На семинаре мы обсудили случайные величины с тонкими хвостами распределения.

Определение 1. *Случайная величина X имеет распределение с тонким хвостом ($X \in LT$), если*

$$\mathbb{P}\left(|X| \geq t\right) \leq 2\exp\{-t/K_1\}.$$

Замечание 1. *Автор называет такие случайные величины субэкспоненциальными, что идет вразрез с устоявшимся и общепринятым определением субэкспоненциального распределения (введено Чистяковым В. П. в 1964 году в работе "Теорема о суммах независимых положительных случайных величин и ее приложения к ветвящимся случайным процессам").*

Изучение случайных величин из класса LT обусловлено важностью для различных приложений неравенств концентрации для евклидовой нормы гауссовских векторов.

Мы подробно обсудили теорему о характерных свойствах класса LT (аналогичную теореме о свойствах субгауссовских случайных величин).

Теорема 1. *Let X be a random variable. Then the following properties are equivalent.*

1. *The tails of X satisfy*

   $$\mathbb{P}\left(|X| \geq t\right) \leq 2\exp\{-t/K_1\}$$

   for all $t \geq 0$.

2. *The moments of X satisfy*

   $$\|X\|_{L^p} = (\mathbb{E}|X|^p)^{1/p} \leq K_2 p \quad \text{for all } p \geq 1.$$

3. *The MGF of $|X|$ satisfies*

   $$\mathbb{E}\exp(\lambda|X|) \leq \exp\left(K_3\lambda\right)$$

   for all λ such that $0 \leq \lambda \leq 1/K_3$.

4. *The MGF of $|X|$ is bounded at some point, namely*

   $$\mathbb{E}\exp\left(|X|/K_4\right) \leq 2$$

   Moreover, if $\mathbb{E}X = 0$ then properties $(i) - (iv)$ are also equivalent to the following property.

5. The MGF of X satisfies

$$\mathbb{E}\exp(\lambda X) \le \exp\left(K_5^2\lambda^2\right)$$

for all λ such that $|\lambda| \le 1/K_5$

The parameters $K_i > 0$ appearing in these properties differ from each other by at most an absolute constant factor.

Замечание 2. *Еще раз обсудили абсолютную константу $C_* > 1$, которая связывает значения $K_1 - K_5$. Предположение о том, что она не зависит от распределения X выглядит правдоподобным.*

Доказательство теоремы не обсуждали, т.к. оно аналогично доказательству теоремы для субгауссовских случайных величин.

Далее, мы ввели норму Орлича

$$\|X\|_{\psi_1} = \inf\{t > 0 : \mathbb{E}\exp(|X|/t) \le 2\}$$

для случайных величин из класса LT и обсудили следующие утверждения об их связи с субгауссовскими:

Утверждение 1. *(LT is subgaussian squared) A random variable X is subgaussian if and only if $X^2 \in LT$. Moreover,*

$$\left\|X^2\right\|_{\psi_1} = \|X\|_{\psi_2}^2$$

Утверждение 2. *(The product of subgaussians is LT) Let X and Y be subgaussian random variables. Then $XY \in LT$. Moreover,*

$$\|XY\|_{\psi_1} \le \|X\|_{\psi_2}\|Y\|_{\psi_2}$$

Затем, мы рассмотрели неравенства концентрации для сумм независимых случайных величин из класса LT — неравенства Бернштейна.

Теорема 2. *(Bernstein's inequality) Let $X_1, \dots, X_N \in LT$ be independent mean-zero random variables. Then, for every $t \ge 0$, we have*

$$\mathbb{P}\left\{\left|\sum_{i=1}^N X_i\right| \ge t\right\} \le 2\exp\left(-c\min\left(\frac{t^2}{\sum_{i=1}^N \|X_i\|_{\psi_1}^2}, \frac{t}{\max_i \|X_i\|_{\psi_1}}\right)\right)$$

where $c > 0$ is an absolute constant.

В частном случае $a_i = 1/N$ получаем неравенство Бернштейна для средних.

Следствие 1. *(Bernstein's inequality) Let $X_1, \dots, X_N \in LT$ be independent mean-zero random variables. Then, for every $t \geq 0$, we have*

$$\mathbb{P}\left\{\left|\frac{1}{N}\sum_{i=1}^{N} X_i\right| \geq t\right\} \leq 2\exp\left(-c\min\left(\frac{t^2}{K^2}, \frac{t}{K}\right)N\right)$$

where $K = \max_i \|X_i\|_{\psi_1}$

Оставшаяся часть семинара была посвящена решению задачи 2.7.3.

Задача. *Consider the class of distributions whose tail decay is of the type* $\exp(-ct^{\alpha})$ *or faster. Here $\alpha = 2$ corresponds to subgaussian distributions and $\alpha = 1$ to LT. State and prove a version of Theorem 1 for such distributions.*