高维概率

High-Dimensional Probability

七、随机向量Concentration

@滕佳烨

- **上节课说了啥**

1. **什么是次高斯随机向量？---用低维定义高维**
   和任意普通向量的内积都是次高斯

2. **一些随机向量的例子**
   Gaussian, Spherical, Symmetric Bernoulli,
   Coordinate, Isotropic Convex
   Isotropic? Sub-Gaussian? Independent coordinate?

- **这节课要说啥**

各种情况下随机向量的Concentration

- **各种情况？**

Coordinate-wise RV
Sub-Gaussian RV

- **Concentration?**

一维变量 -> $P(|X| \geq t) \leq u$
高维变量 -> $P(\|X\| \geq t) \leq u$

- Coordinate-wise concentration(1)

Theorem 3.1.1 (Concentration of the norm). *Let $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ be a random vector with independent, sub-gaussian coordinates X_i that satisfy $\mathbb{E}\, X_i^2 = 1$. Then*

$$\left\| \|X\|_2 - \sqrt{n} \right\|_{\psi_2} \le CK^2,$$

where $K = \max_i \|X_i\|_{\psi_2}$ and C is an absolute constant.[1]

- Coordinate!!!

注意到Coordinate-wise次高斯是次高斯随机向量的**子集**

我们之前提到的**高斯≈球壳**就是它的一个特例嘛！

- Coordinate-wise concentration(1)

Theorem 3.1.1 (Concentration of the norm). *Let $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ be a random vector with independent, sub-gaussian coordinates X_i that satisfy $\mathbb{E} X_i^2 = 1$. Then*

$$\Big\| \|X\|_2 - \sqrt{n} \Big\|_{\psi_2} \le C K^2,$$

where $K = \max_i \|X_i\|_{\psi_2}$ and C is an absolute constant.[1]

- Hint

1. 由于X_i是次高斯， X_i^2就是次指数，就能找到$\|X\|^2$的bound
2. 如果$z \ge 0$，$|z-1| \ge \delta \to |z^2 - 1| \ge \max(\delta, \delta^2)$，逆否

- Coordinate-wise concentration(2)

Theorem 6.3.2 (Concentration of random vectors). *Let B be an $m \times n$ matrix, and let $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ be a random vector with independent, mean zero, unit variance, sub-gaussian coordinates. Then*

$$\Big\| \|BX\|_2 - \|B\|_F \Big\|_{\psi_2} \le CK^2 \|B\|,$$

where $K = \max_i \|X_i\|_{\psi_2}$.

- 若B=I，就是之前的定理！
- 仍然是Coordinate-wise
- ***那么对于一个一般的次高斯随机向量，是否也有类似的结果呢？很遗憾，**没有。但是我们有下一页给出的定理**

- Sub-Gaussian Concentration*

Exercise 6.3.5 (Tails of sub-gaussian random vectors). ☕☕ Let B be an $m \times n$ matrix, and let X be a mean zero, sub-gaussian random vector in $\mathbb{R}^n$ with $\|X\|_{\psi_2} \le K$. Prove that for any $t \ge 0$, we have

$$\mathbb{P}\left\{\|BX\|_2 \ge CK\|B\|_F + t\right\} \le \exp\left(-\frac{ct^2}{K^2\|B\|^2}\right).$$

- 注意到这里只有一个下界结果，与之对比的，coordinate 还有对应的上界结果

1. Coordinate-wise sub-Gaussian向量有Concentration结果

$$\|X\| \approx \sqrt{n}$$

$$\|BX\| \approx \|B\|_F$$

2. 普通的次高斯向量只有下界没上界

谢谢！

@滕佳烨