High dimensional probability 04/13/2022

Uniform distribution on the unit Euclidean sphere in $\mathbb{R}^n$

Bowen Zhou *bowen_zz@sjtu.edu.cn*

Consider a random vector X uniformly distributed on the unit Euclidean sphere in $\mathbb{R}^n$ with center at the origin and radius 1, denote as $X \sim Unif(S^{n-1})$.

Moments of X

(a) $\mathbb{E}X = 0$, $Cov(X) = \frac{1}{n} I_n$;

(b) $\mathbb{E}X_i^{2k-1} = 0$, $\mathbb{E}X_i^{2k} = \frac{4k-1}{n+4k-2} \frac{4k-3}{n+4k-4} \cdots \frac{1}{n}$.

Proof.

(a) By symmetric property, $\mathbb{E}X = 0$. $Cov(X) = \frac{1}{n} I_n \Leftrightarrow \mathbb{E}(x^T X)^2 = \|x\|^2/n$ for any $x \in \mathbb{R}^n$. By rotation invariance, for all $x, y \in \mathbb{R}$ with $\|x\| = \|y\|$, we have $\mathbb{E}(x^T X)^2 = (y^T X)^2$, so:

$$\mathbb{E}(x^T X)^2 = \frac{\|x\|^2}{n} \sum_{i=1}^{n} \mathbb{E}(e_i^T X)^2 = \frac{\|x\|^2}{n} \sum_{i=1}^{n} \mathbb{E}X_i^2 = \frac{\|x\|^2}{n}.$$

(b) To compute higher moments of X, we need more tools rather than symmetric property. Indeed, the uniform distribution on the sphere is quite similar to the normal distribution.

Lemma 1. *Let $g \sim N(0, I_n)$, and represent is as $g = r\theta$ where $r = \|g\|_2, \theta = g/\|g\|_2$, then*

(a) r and θ are independent random variables;
(b) $\theta \sim Unif(S^{n-1})$.

By the lemma, we have $X \overset{d}{=} g/\|g\|_2$. Now let $X = (X_1, X_2, \cdots, X_n)$ and $g = (g_1, g_2, \cdots, g_n)$, then

$$\mathbb{E}g_i^m = \mathbb{E}\|g\|_2^m \frac{g_i^m}{\|g\|_2^m} = \mathbb{E}\|g\|_2^m \mathbb{E}\frac{g_i^m}{\|g\|_2^m},$$

$\|g\|_2^2 \sim \chi^2(n)$ and $\mathbb{E}g_i^{2k-1} = 0$ so $\mathbb{E}X_i^{2k-1} = 0$. When $m = 2k$, the moments of normal and chi distribution tells us

$$\mathbb{E}g_i^{2k} = (4k-1)!!,\ \mathbb{E}\|g\|_2^{2k} = [n+4k-2][n+4k-4]\cdots n,$$

so $\mathbb{E}X_i^{2k} = \frac{4k-1}{n+4k-2} \frac{4k-3}{n+4k-4} \cdots \frac{1}{n}$.

□

Similarly, any moments $\mathbb{E}X_1^{k_1} X_2^{k_2} \cdots X_n^{k_n}$ can be calculated.

Concentration

Now we consider the concentration inequality of this kind of sphere distribution, actually, it's a sub-gaussian random vector. By standardization, we just consider $X \sim Unif(\sqrt{n} S^{n-1})$ of which the covariance matrix is an identity matrix.

Proof Idea $X = \sqrt{n} \frac{g}{\|g\|_2}$ so we just need to establish the concentration of $\|g\|_2$. Actually, if a random vector Y has independent sub-gaussian cordinates, then the norm of Y is also sub-gaussian. So it is easy to get the conclusion. □