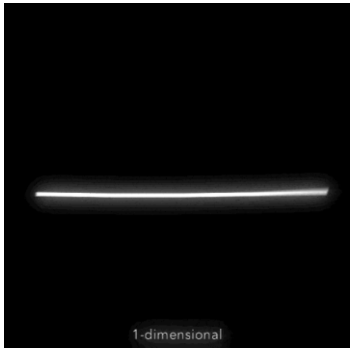
课程主页：http://staff.ustc.edu.cn/~ynyang/vector

第二讲 球对称分布

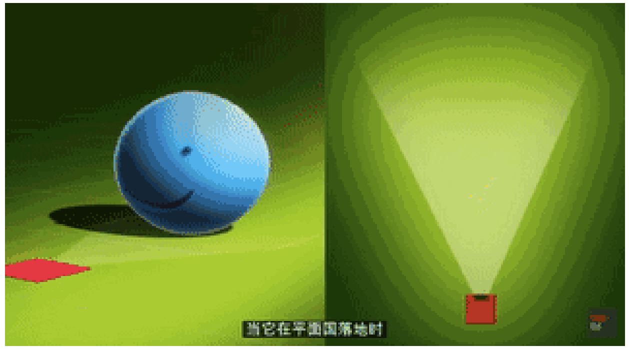
2024.2.28

球面均匀分布是球面上唯一正交变换下不变的分布

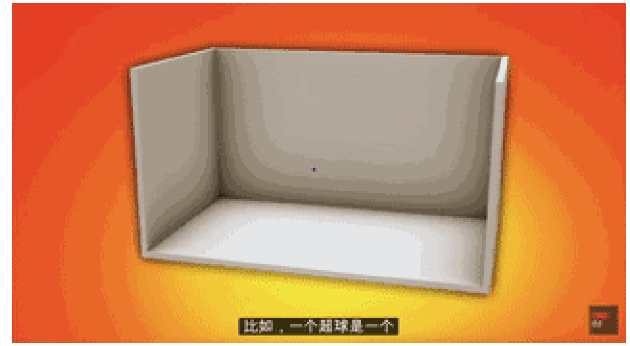
Recap



展开立方体



穿越2维平面的三维球



穿越3维空间的四维球

https://www.zhihu.com/zvideo/1409650387691003904

第一讲定理1给出了n维欧氏空间中单位球面S^{n-1}上的均匀分布的边际概率密度，在某种意义上描述了球面的几何性质。边际概率密度或其它概率运算最好能给出显式数学表达（如定理1），也可以用蒙特卡洛方法模拟计算，两者互相印证。

关于$n=5$降维打击： $\mathbf{u} \sim U(S^4) \subset R^5 \Rightarrow \mathbf{u}_{[1:3]} \sim U(B^3)$

我们隐含假设了星体都是球形的，生物生活在球表面。

- 引力本质是空间的弯曲（球面，非欧几何）
- 球形：旋转不变，各向同性

> 摘自网络杂志“Aeon”科普文章“Radical dimensions” by Margaret Wertheim:
> https://aeon.co/essays/how-many-dimensions-are-there-and-what-do-they-do-to-reality
>
> 弦理论还有一个值得关注的发展。1999 年，丽莎•兰德尔（Lisa Randall，第一位在哈佛大学获得终身教职的理论物理学家）和拉曼•桑德拉姆（Raman Sundrum，一位印度裔美国粒子理论家）提出，在广义相对论所描述的宇宙学尺度上，可能存在一个额外的维度。根据他们的“brane”理论——“brane”是“膜”的缩写——我们通常所指的宇宙可能嵌在一个更大的五维空间，一种超级宇宙中。在这个超级空间中，我们的宇宙可能只是众多平行宇宙中的一个，它们每一个都是一个独立的四维气泡，位于更广阔的五维空间里。

我们知道几何概率常常会出现悖论（比如Betrand paradox），这与均匀或随机的理解或定义的不同有关，也与球面上均匀分布没有概率密度有关。我们将利用球对称分布（比如标准正态分布）给出球面均匀分布随机向量的随机表示（并将以此为工具证明第一讲定理1）。

本次课：首先得到正态分布的以χ分布和球面均匀分布的正交随机表示刻画；将χ分布推广，得到一般的正交不变分布的刻画；在仿射变换下得到椭球对称分布, 包括多元(椭球)正态。

惯例:
虚线框内容可忽略;
附录内容课忽略;

符号表示:
向量/随机向量：黑正体小写字母，$\mathbf{x}, \boldsymbol{\theta}$
变量/随机变量：斜体小写字母，x, θ （有时也用大写字母)
矩阵：大写字母，X

多元标准正态与χ、χ^2分布

标准正态

多元标准正态分布: $\mathbf{x}=(x_1,\ldots,x_n)^{\mathrm{T}} \sim N_n(0,I_n) \Leftrightarrow x_1,\ldots,x_n \ \ iid \sim N(0,1)$

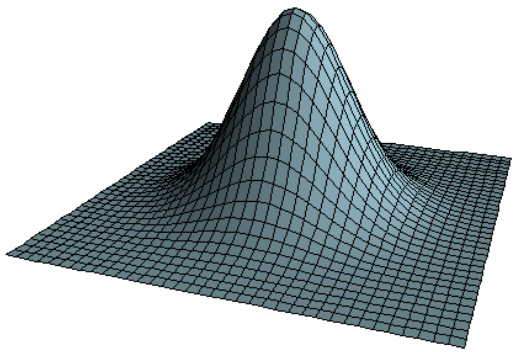
$\Leftrightarrow$ 概率密度函数: $p(\mathbf{x})=\dfrac{1}{(2\pi)^{n/2}}\exp\left(-\dfrac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{x}\right)$

球对称

$\mathbf{x}\sim N_n(0,I_n)$的密度仅与模长$\|\mathbf{x}\|$有关,

对任何$\forall$正交阵H，$H\mathbf{x}\overset{d}{=}\mathbf{x}$,即$\mathbf{x}$的分布

是正交旋转不变的，或球对称的。



卡方分布

若$\mathbf{x}\sim N_n(0,I_n)$,称$\|\mathbf{x}\|^2$服从自由度为$n$的卡方分布,$\|\mathbf{x}\|^2\sim\chi_n^2$,

$\|\mathbf{x}\|$服从χ_n分布。

R^n球面面积：$A=\dfrac{2\pi^{n/2}}{\Gamma(n/2)}r^{n-1}$,

R^n球的体积：$V=Ar/n=\dfrac{2\pi^{n/2}}{n\Gamma(n/2)}r^{n}$.

下面求χ_n的分布（由此立得χ_n^2的分布）

解：$\mathbf{x}$的密度$f(\mathbf{x})=\exp(-\|\mathbf{x}\|^2/2)/(2\pi)^{n/2}$.

半径r到$r+\Delta r$的球壳体积为$\Delta V=|S^{n-1}(r)|\Delta r=\dfrac{2\pi^{n/2}}{\Gamma(n/2)}r^{n-1}\Delta r$,

其中$|S^{n-1}(r)|=\dfrac{2\pi^{n/2}}{\Gamma(n/2)}r^{n-1}$为球面面积,所以球壳上的概率质量为

$$f(\mathbf{x})\Delta V=f(\mathbf{x})\times\frac{2\pi^{n/2}}{\Gamma(n/2)}r^{n-1}\Delta r=\frac{1}{2^{n/2-1}\Gamma(n/2)}r^{n-1}e^{-r^2/2}\Delta r=p_{\chi_n}(r)\Delta r$$

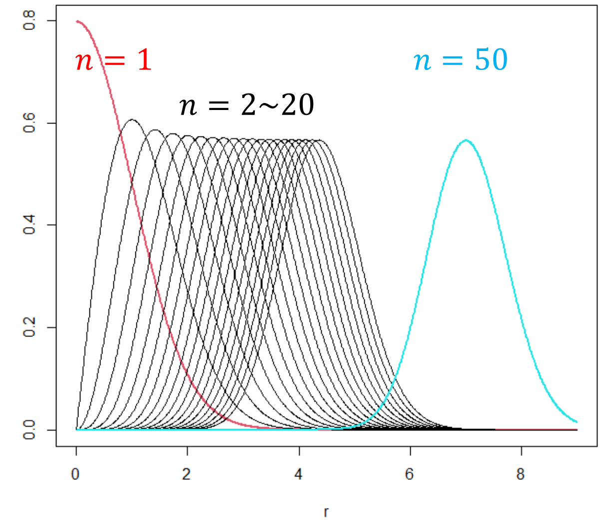
所以$p_{\chi_n}=\dfrac{1}{2^{n/2-1}\Gamma(n/2)}r^{n-1}e^{-r^2/2}$。

注：也可用球坐标方法（参见定理2）

χ_n的概率密度：

$$p_{\chi_n}(r) = \frac{1}{2^{n/2-1}\Gamma(n/2)} r^{n-1} e^{-r^2/2}, r > 0$$

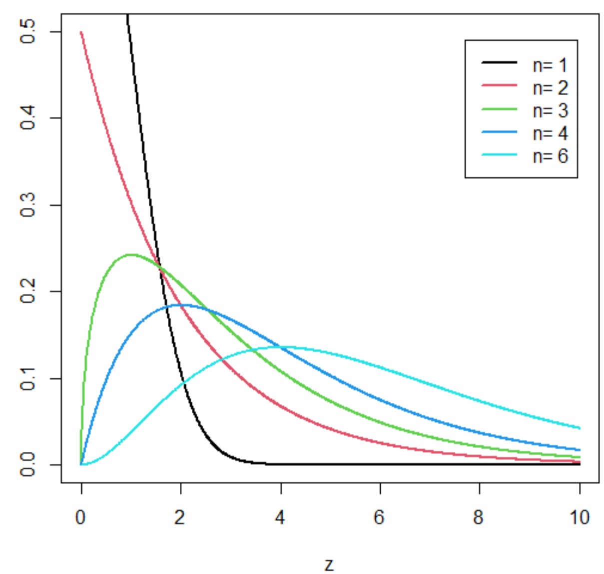
- $n = 1$时，密度单调下降；
- $n \geq 2$时，单峰分布，众数mode $= \sqrt{n-1}$；
- 当n较大，密度形状几乎相同，n变化时密度几乎平移不变（方差近似相同）。后面将证明$\chi_n \approx N(\sqrt{n-1}, 1/2)$。



χ_n^2的概率密度：

$$p_{\chi_n^2}(z) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} e^{-z/2}, \ z > 0。$$

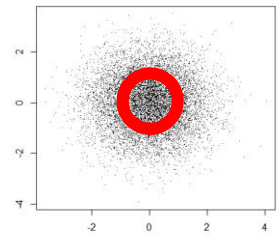
- $n = 1, 2$时，密度单调下降；
- $n = 2$时，χ_2^2是指数分布（参数1/2）；
- $n \geq 3$时，单峰分布，众数mode $= n - 2$。



众数
mode

$$\mathrm{mode}(\chi_n)=\sqrt{n-1},\ n\ge 2;$$
$$\mathrm{mode}(\chi_n^2)=n-2,\quad n\ge 3。$$

注：$\mathbf{x}\sim N_n(0,I_n)$，$n>1$时，考虑到原点的距离，$\mathbf{x}$最大可能落在半径为$\sqrt{n-1}$的球面上。对于二元标准正态，单位圆上数据点最多。



均值方差

- $E(\chi_n)=\sqrt{2}\Gamma(n/2)/\Gamma((n-1)/2)$，　$\mathrm{var}(\chi_n)=(n-1)-(E\chi_n)^2$
  $n\to\infty$时，$E(\chi_n)\approx\sqrt{n-1}$, $\mathrm{var}(\chi_n)\approx 1/2.$
- $E(\chi_n^2)=n, \mathrm{var}(\chi_n^2)=2n$

记号说明：
- $E(\chi_n)$表示χ_n分布的期望，完整的表达为“$E(X)$，$X\sim\chi_n$”
- “$\chi_n^2\approx N(n,2n)$”即“若$X\sim\chi_n^2$，则近似地$X\sim N(n,2n)$”

Gamma与χ^2分布

$$\frac{1}{2}\chi_n^2 \overset{d}{=} \text{Gamma}(n/2)\text{分布，其中Gamma}(\alpha)\text{分布}: p(t)=\frac{1}{\Gamma(\alpha)}t^{\alpha-1}e^{-t},$$

$$\chi_n^2 \overset{d}{=} \text{Gamma}(n/2,1/2)\text{分布，其中Gamma}(\alpha,\lambda): p(t)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}t^{\alpha-1}e^{-\lambda t}$$

分布近似

$n \longrightarrow \infty$ 时， $\chi_n^2 \approx N(n,2n)$； $\chi_n \approx N(\sqrt{n-1},1/2)$，

其中“$\approx$”表示两端的分布近似相同。

n个独立标准正态随机变量的平方和$\sum_{i=1}^n z_i^2 \sim \chi_n^2$，其中$z_i$ iid$\sim N(0,1)$,

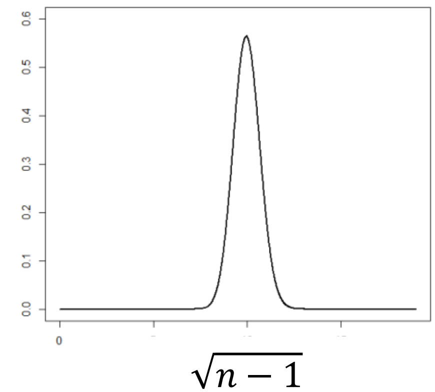
由中心极限定理

$$\frac{\sum_{i=1}^n z_i^2-n}{\sqrt{2n}} \overset{d}{\longrightarrow} N(0,1),\quad n\longrightarrow\infty$$

换言之，$\chi_n^2 \approx N(n,2n)$。 由delta方法, $\chi_n \approx N(\sqrt{n-1},1/2)$。

> 注： $\chi_n \approx N(\sqrt{n-1}, 1/2)$ 这一近似令人吃惊：
> 若$\mathbf{x} \sim N_n(0, I_n)$, 则其模长$\|\mathbf{x}\|$近似地有相同形状的分布，形状与维数无关：$\|\mathbf{x}\| \approx N\left(\sqrt{n-1}, \frac{1}{2}\right)$.

$\|\mathbf{x}\| \approx N(\sqrt{n-1}, 1/2)$



由上述近似正态结果可知，若$\mathbf{x} \sim N_n(0, I_n)$，则

$$P(\sqrt{n-1} - 1.38 \le \|\mathbf{x}\| \le \sqrt{n-1} + 1.38) \approx 0.95,$$

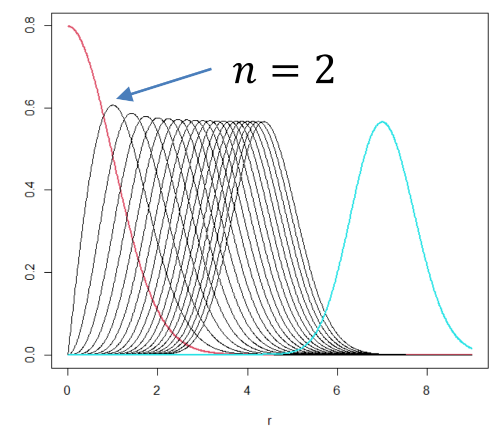
其中$1.38 = 1.96/\sqrt{2}$。

所以

- n较大时，$N_n(0, I_n)$的约95%的概率质量集中在半径$\sqrt{n-1}$，厚度为2.76的球台上。
- 更令人吃惊的是，即使$n=2$，χ_2已经接近正态：

$$P(\sqrt{2-1} - 1.38 \le \|\mathbf{x}\| \le \sqrt{2-1} + 1.38) = 0.941,$$

（参见χ_n分布概率密度图）



球对称分布

$\overset{d}{=}$ 或~表示等分布

球对称分布

定义(球对称分布). 若对任何正交矩阵H, $H\mathbf{x} \overset{d}{=} \mathbf{x}$, 则称$\mathbf{x}$服从球对称分布（或正交不变分布）。

其它名称：球对称/球等高/正交不变/旋转不变分布，Spherically symmetric distribution, spherically contoured distribution, orthogonally invariant distribution, rotationally invariant distribution

命题1. 若$\mathbf{x}$有概率密度函数$f(\mathbf{x})$，且$f(\mathbf{x})$仅依赖于$\|\mathbf{x}\|$, 即存在一元函数h使得

$$f(\mathbf{x}) = h(\|\mathbf{x}\|)$$

则$\mathbf{x}$服从球对称分布。

球对称分布的例子：

- R^1中的球对称分布：0对称分布 (1阶正交矩阵只有1和−1)。
- $N_n(0, \sigma^2 I_n)$， $f(\mathbf{x}) = \frac{1)}{(2\pi\sigma^2)^{n/2}} e^{-\|\mathbf{x}\|^2/2\sigma^2}$
- $U(B^n(r))$, $f(\mathbf{x}) = \frac{1}{|B^n(R)|} = \frac{\Gamma(1+n/2)}{\pi^{n/2} r^n}, \|\mathbf{x}\| \le r$
- $U(S^{n-1})$（无密度，命题2 ）

其它球对称分布的例子：

- n元t分布(自由度k)：

$$f(\mathbf{x})=\frac{\Gamma((n+k)/2)}{(k\pi)^{n/2}\Gamma(k/2)}\left(1+\frac{\|\mathbf{x}\|^2}{k}\right)^{-(n+k)/2},\quad \mathbf{x}\in R^n$$

($n=1$时为一元student's t分布; $k=1$时为n元Cauchy分布)

- n元Cauchy分布：

$$f(\mathbf{x})=\frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}\left(1+\|\mathbf{x}\|^2\right)^{-(n+1)/2}$$

同分布等号

同分布操作需要小心，比如$x\overset{d}{=}y$蕴含$x+c\overset{d}{=}y+c$ （常数c),
但未必蕴含$x+z\overset{d}{=}y+z$ （随机变量z ）

引理1：若随机向量$\mathbf{x}\overset{d}{=}\mathbf{y}$， 则$g(\mathbf{x})\overset{d}{=}g(\mathbf{y})$, 其中$g$是任一（可测）向量函数。

特别地，$x_i\overset{d}{=}y_i$，$x_ix_j\overset{d}{=}y_iy_j$.

若**x**服从球对称分布，直观上**x**应等可能地/均匀地分布在各个方向上，我们将在定理**2**给出严格证明。首先看一个例子：

例1（一元球对称分布）.

（1）若$P(x=0)=0$, 且x服从关于0点的一元对称分布, 即 $x \overset{d}{=} -x$, 则 $u=x/|x|$与$|x|$独立, 且$u \sim U\{\pm 1\}$。

（2）x服从关于0点的一元对称分布$\Leftrightarrow$ 对任何$u \sim U\{\pm 1\}$，$u \perp |x|$, 有 $x \overset{d}{=} |x|u$

证明：（1）.由对称性$P(u=1)=P(x>0)=\frac{1}{2}=P(x<0)=P(u=-1)$, 另外$P(|x| \le t, u=1)=P(0<x\le t)=\frac{1}{2}P(|x|\le t)=P(u=-1)\times P(|x|\le t)$，所以 u与$|x|$独立。

例2. 假设$x, y\ iid \sim N(0,1)$, $u=y/|y| \sim U\{\pm 1\}$，由例1，x与x/u同分布，则$\frac{x}{y}$与$\frac{x}{|y|}$同分布，这是因为

$$\frac{x}{y}=\frac{x}{|y|u}=\frac{x/u}{|y|}\overset{d}{=}\frac{x}{|y|}$$

该共同分布为t_1或Cauchy分布。

命题2：若$\mathbf{x}$服从球对称分布， 则$\mathbf{x}$各个分量同分布，且

$$E(\mathbf{x}) = \mathbf{0}, var(\mathbf{x}) = \sigma^2 I_n$$

证明：因为置换变换是正交变换，所以 $\mathbf{x}$各个分量同分布，方差相同；因为反射(reflection: $x \to -x$)是正交变换，我们有 $x_i \overset{d}{=} -x_i$， $x_i x_j \overset{d}{=} -x_i x_j$，从而$E(x_i) = 0, cov(x_i, x_j) = 0.$

命题3. $\mathbf{x}$服从球对称分布 $\Leftrightarrow \mathbf{t}^\top \mathbf{x} \overset{d}{=} x_1, \forall\, \mathbf{t} \in S^{n-1}$。

证明：-($\Rightarrow$)任取正交矩阵H,其第一行为$\mathbf{t}$, 则 $H\mathbf{x} \overset{d}{=} \mathbf{x}$，特别地两边的第一分量等分布，即$\mathbf{t}^\top \mathbf{x} \overset{d}{=} x_1$。

($\Leftarrow$) 对任何$\mathbf{s} \in R^n$, 记$\mathbf{t} = \mathbf{s}/||\mathbf{s}|| \in S^{n-1}$， $H^\top \mathbf{t} \in S^{n-1}$, 所以

$$x_1 \overset{d}{=} \mathbf{t}^\top \mathbf{x}，x_1 \overset{d}{=} (H^\top \mathbf{t})^\top \mathbf{x} = \mathbf{t}^\top H\mathbf{x},$$

所以

$$\mathbf{t}^\top \mathbf{x} \overset{d}{=} \mathbf{t}^\top H\mathbf{x}$$

两边同乘常数$||\mathbf{s}||$还是同分布的，即$\big|\,|\mathbf{s}|\,\big|\mathbf{t}^\top \mathbf{x} \overset{d}{=} ||\mathbf{s}||\mathbf{t}^\top H\mathbf{x}$，即$\mathbf{s}^\top \mathbf{x} \overset{d}{=} \mathbf{s}^\top H\mathbf{x}$，其矩母函数或特征函数也相同：$E\exp(\mathbf{s}^\top \mathbf{x}) = E\exp(\mathbf{s}^\top H\mathbf{x})$或$E\exp(i\mathbf{s}^\top \mathbf{x}) = E\exp(i\mathbf{s}^\top H\mathbf{x})$，所以$\mathbf{x} \overset{d}{=} H\mathbf{x}$。

命题4. 球面均匀分布是球对称分布，即若$\mathbf{u} \sim U(S^{n-1})$，则对任何正交矩阵$H$，$H\mathbf{u} \sim U(S^{n-1})$.

证明：对任何$S_0 \subset S^{n-1}$和正交矩阵H, $|H^\top S_0| = |S_0|$,则

$P(H\mathbf{u} \in S_0) = P(\mathbf{u} \in H^\top S_0) = |H^\top S_0| / |S^{n-1}| = |S_0| / |S^{n-1}| = P(\mathbf{u} \in S_0)$。

定理1. 随机向量$\mathbf{z} \in S^{n-1}$是球对称的 $\Leftrightarrow \mathbf{z} \sim U(S^{n-1})$

证明：($\Rightarrow$) 取$\mathbf{u} \sim U(S^{n-1})$, $\mathbf{u}$与$\mathbf{z}$独立，下面证明 $\mathbf{z}$与 $\mathbf{u}$同分布。

因为$\mathbf{u}$是球对称的，由命题3， $\mathbf{z}^\top \mathbf{u} | \mathbf{z} \overset{d}{=} u_1 | \mathbf{z} \overset{d}{=} u_1$

（最后一个等号是因为$\mathbf{z}$与u_1独立），所以$\mathbf{z}^\top \mathbf{u} | \mathbf{z} \overset{d}{=} u_1$，即给定$\mathbf{z}$时 $\mathbf{z}^\top \mathbf{u}$的条件分布与条件$\mathbf{z}$无关，所以$\mathbf{z}^\top \mathbf{u} \overset{d}{=} u_1$。

同理，因为$\mathbf{z}$是球对称的， $\mathbf{z}^\top \mathbf{u}|\mathbf{u} \overset{d}{=} z_1|\mathbf{u} \overset{d}{=} z_1 \Rightarrow \mathbf{z}^\top \mathbf{u} \overset{d}{=} z_1$。

所以 $z_1 \overset{d}{=} u_1$。

注：$z_1 \overset{d}{=} u_1 \Rightarrow z_i \overset{d}{=} u_i, i = 1, \ldots, n$,
但这并不蕴含$\mathbf{z} \overset{d}{=} \mathbf{u}$

对任何常数向量 $\mathbf{t} \in S^{n-1}$，因为$\mathbf{u}$， $\mathbf{z}$都是球对称的，

$$\mathbf{t}^\top \mathbf{u} \overset{d}{=} u_1,\ \mathbf{t}^\top \mathbf{z} \overset{d}{=} z_1 \Rightarrow \mathbf{t}^\top \mathbf{z} \overset{d}{=} \mathbf{t}^\top \mathbf{u}$$

所以$E\exp(\mathbf{t}^\top \mathbf{z}) = E\exp(\mathbf{t}^\top \mathbf{u})$, $\mathbf{u}$和$\mathbf{z}$有相同的矩母函数，

从而 $\mathbf{z}$与 $\mathbf{u}$同分布，即$\mathbf{z} \sim U(S^{n-1})$。证毕。

定理1说明$U(S^{n-1})$是单位球面上唯一的正交不变分布。

对于$\mathbf{x}$具有一般球对称分布的情形，我们也有类似于第一讲命题1的结论。因为球坐标的角度一般不是我们感兴趣的，所以结论的表述将只使用模长$r=\|\mathbf{x}\|$和$\mathbf{x}$的单位化球面向量$\mathbf{u}=\frac{\mathbf{x}}{\|\mathbf{x}\|}\in S^{n-1}$.

定理2：假设$\mathbf{x}$服从球对称分布（概率密度函数为f），即存在一元函数h,使得 $f(\mathbf{x})=h(\|\mathbf{x}\|)$,则

$$\mathbf{x}/\|\mathbf{x}\|\sim U(S^{n-1}),\ \ \mathbf{x}/\|\mathbf{x}\| \text{ 与 } \|\mathbf{x}\| \text{独立,}$$

且$\|\mathbf{x}\|$的概率密度为 $p(r)=\dfrac{2\pi^{n/2}}{\Gamma(n/2)}h(r)r^{n-1}, r>0.$

第一讲引理1：球坐标变换$\mathbf{x}\to(r,\boldsymbol{\theta})=(r,\theta_1,\ldots,\theta_{n-1})$，$r\geq 0$，$0\leq\theta_1,\ldots,\theta_{n-2}\leq\pi$, $0\leq\theta_{n-1}<2\pi$，

$$\begin{aligned}
x_1 &= r\cos(\theta_1)=ru_1\\
x_2 &= r\sin(\theta_1)\cos(\theta_2)=ru_2\\
&\ldots\\
x_{n-1} &= r\sin(\theta_1)\cdots\sin(\theta_{n-2})\ \cos(\theta_{n-1})=ru_{n-1}\\
x_n &= r\sin(\theta_1)\cdots\sin(\theta_{n-2})\ \sin(\theta_{n-1})=ru_n
\end{aligned}$$

变换的 Jacobi：

$$J(\mathbf{x}\to(r,\boldsymbol{\theta}))=r^{n-1}\sin^{n-2}(\theta_1)\sin^{n-3}(\theta_2)\cdots\sin(\theta_{n-2}) .$$

定理2的证明：记$\mathbf{u}=\mathbf{x}/\|\mathbf{x}\|$， $r=\|\mathbf{x}\|$。

应用球坐标变换$\mathbf{x}\to(r,\boldsymbol{\theta})=(r,\theta_1,\ldots,\theta_{n-1})$， $(r,\boldsymbol{\theta})$的概率密度

$$p(r,\boldsymbol{\theta})=h(r)\times J=h(r)\times r^{n-1}\sin^{n-2}(\theta_1)\sin^{n-3}(\theta_2)\cdots\sin(\theta_{n-2})$$

这表明$r,\theta_1,\ldots,\theta_{n-1}$相互独立，而$\mathbf{u}=\mathbf{x}/\|\mathbf{x}\|=\mathbf{x}/r$仅与$\theta's$有关，$\|\mathbf{x}\|=r$,

所以$\|\mathbf{x}\|$ 与$\mathbf{u}=\mathbf{x}/\|\mathbf{x}\|$独立。

注意：只有$\theta_{n-1}\sim U(0,2\pi)$
均匀分布，其它 $\theta_i's$ 不服从均匀分布

因为

$$|S^{n-1}|=\int\sin^{n-2}(\theta_1)\sin^{n-3}(\theta_2)\cdots\sin(\theta_{n-2})\ d\boldsymbol{\theta}=\frac{2\pi^{n/2}}{\Gamma(n/2)}$$

所以

$$p(r,\boldsymbol{\theta})=\left[\frac{2\pi^{n/2}}{\Gamma(n/2)}h(r)r^{n-1}\right]\times\left[\frac{\Gamma(n/2)}{2\pi^{n/2}}\sin^{n-2}(\theta_1)\sin^{n-3}(\theta_2)\cdots\sin(\theta_{n-2})\right]$$

$$\triangleq p(r)\times q(\boldsymbol{\theta})$$

两个方括号内分别是$r,\boldsymbol{\theta}$的概率密度$p(r),q(\boldsymbol{\theta})$ 。

下面证明$\mathbf{u}\sim U(S^{n-1})$：由定理1， $U(S^{n-1})$是球面上唯一正交不变的分布, 我们只需证明对任何正交矩阵H, $H\mathbf{u}\sim\mathbf{u}$ ($H\mathbf{u}$和$\mathbf{u}$同分布）。

由$H\mathbf{x}\sim \mathbf{x} \Rightarrow H\mathbf{x}/\|H\mathbf{x}\|\sim\mathbf{x}/\|\mathbf{x}\|$，又因$\|\mathbf{x}\| = \|H\mathbf{x}\|$

$$\Rightarrow H\mathbf{u} = H\mathbf{x}/\|\mathbf{x}\| = H\mathbf{x}/\|H\mathbf{x}\| \sim \mathbf{x}/\|\mathbf{x}\| = \mathbf{u}$$

所以 $\mathbf{u} = \mathbf{x}/\|\mathbf{x}\| \in S^{n-1}$的分布正交不变，所以$\mathbf{u}\sim U(S^{n-1})$. 证毕。

注：实际上$q(\boldsymbol{\theta})$是球面均匀分布的极坐标下的分布密度, 利用球坐标也可以证明 $\mathbf{u}\sim U(S^{n-1})$：

对任何$S_0\subset S^{n-1}$

$$P(\mathbf{u}\in S_0) = \int_{\mathbf{u}\in S_0} q(\boldsymbol{\theta})d\boldsymbol{\theta}$$

$$= \int_{\mathbf{u}\in S_0} \frac{\Gamma(n/2)}{2\pi^{n/2}} \sin^{n-2}(\theta_1)\sin^{n-3}(\theta_2)\cdots\sin(\theta_{n-2})\ d\boldsymbol{\theta}$$

$$= \frac{\Gamma(n/2)}{2\pi^{n/2}}\int_{\mathbf{u}\in S_0} d_{S^{n-1}}A = \frac{\Gamma(n/2)}{2\pi^{n/2}}|S_0| = |S_0|/|S^{n-1}|$$

单位球面的面积元（第一讲引理1）
$d_{S^{n-1}}A=\sin^{n-2}(\theta_1)\sin^{n-3}(\theta_2)\cdots\sin(\theta_{n-2})\,d\boldsymbol{\theta}$

所以$\mathbf{u}\sim U(S^{n-1})$.

从定理2 的证明容易得到如下刻画（随机表示，同分布表示）

球对称分布的刻画

定理3：$\mathbf{x}$服从球对称分布$\Leftrightarrow$ $\mathbf{x}\overset{d}{=}\mathbf{u}r$，其中$\mathbf{u}$，$r$独立，

$\mathbf{u}\sim U(S^{n-1})$，$r\overset{d}{=}\|\mathbf{x}\|$。

$\mathbf{x}=\mathbf{u}r$
这种方向$\mathbf{u}$乘以长度r的分解叫做“极分解”

证明：（$\Rightarrow$）取$\mathbf{u}=\mathbf{x}/\|\mathbf{x}\|$，$r=\|\mathbf{x}\|$。

（$\Leftarrow$）若$\mathbf{x}\overset{d}{=}\mathbf{u}r$,因为$\mathbf{u}\sim U(S^{n-1})$正交不变，所以对任何正交矩阵$H$，

$H\mathbf{x}\overset{d}{=}H\mathbf{u}r\overset{d}{=}\mathbf{u}r\overset{d}{=}\mathbf{x}$，即$\mathbf{x}$分布正交不变，$\mathbf{x}$服从球对称分布。

注：模长分布唯一决定了球对称分布，具体关系如下：

已知$\mathbf{x}\sim f(\mathbf{x})=h(||\mathbf{x}||)\Rightarrow||\mathbf{x}||\sim p(r)=\dfrac{2\pi^{n/2}}{\Gamma(n/2)}h(r)r^{n-1}$

已知$||\mathbf{x}||\sim p(r)\Rightarrow \mathbf{x}\sim f(\mathbf{x})=h(||\mathbf{x}||)=\dfrac{\Gamma\left(\frac{n}{2}\right)p(||\mathbf{x}||)}{2\pi^{n/2}||\mathbf{x}||^{n-1}}$

多元标准正态是定理2，3的特殊情况，鉴于其重要性，我们把它单独列为定理 4。

定理4.

(1) $\mathbf{x} \sim N_n(0, I_n) \Leftrightarrow \mathbf{x}/\|\mathbf{x}\| \sim U(S^{n-1})$，$\mathbf{x}/\|\mathbf{x}\| \perp\!\!\!\perp \|\mathbf{x}\|$，$\|\mathbf{x}\| \sim \chi_n$.

(2) $\mathbf{x} \sim N_n(0, I_n) \Leftrightarrow \mathbf{x} \overset{d}{=} \mathbf{u}r$, 其中 $\mathbf{u} \sim U(S^{n-1}), r \sim \chi_n$，$\mathbf{u} \perp\!\!\!\perp r$.

正态的随机表示（2）形式地写成：

$$N_n(0, I_n) = U(S^{n-1}) \times \chi_n$$

注：

可利用定理4结论（2）和球面均匀分布生成一元正态随机数（例3）。

可利用定理4结论（1）和球对称分布生成球面均匀随机向量

例3. （二元标准正态的随机表示/极分解及其应用)

$n=2$时，$\mathbf{x}\sim N_2(0,I_2)$有表示$\mathbf{x}\overset{d}{=}\mathbf{u}R$，其中 $\mathbf{u}=(u_1,u_2\ )^\top\sim U(S^1)$单位圆周上的均匀分布，$R\sim\chi_2$具有概率密度 $p_{\chi_2}(r)=r\exp(-r^2/2)$。
该表示可用来从$\mathbf{u}$和R生成正态随机数，如何生成$\mathbf{u}$和R？

- 取 $\theta_1\sim(0,2\pi)$, 则 $\mathbf{u}=(u_1,u_2\ )^\top=(\cos(\theta_1),\sin(\theta_1))^\top\sim U(S^1)$
- 而χ_2的累积分布函数为$F(r)=1-\exp(-r^2/2)$, 则$F(R),\bar{F}(R)$ $\sim U(0,1)$, 记$\xi\triangleq\bar{F}(R)=\exp\left(-\frac{R^2}{2}\right)\sim U(0,1)$ ， 故$R=\sqrt{-2\log(\xi)}$，其中$\xi\sim U(0,1)$ 。

Box-Muller方法: 产生$\mathbf{x}=(x_1,x_2\ )^\top\sim N_2(0,I_2)$, 即$x_1$, x_2 iid $N(0,1)$

- 产生$\theta_1\sim U(0,2\pi)$，$\xi\sim U(0,1)$，独立;
- $R=\sqrt{-2\log(\xi)}$， $\mathbf{u}=(\cos(\theta_1),\sin(\theta_1))^\top$;
- $\mathbf{x}=R\mathbf{u}$.

利用Box-Muller方法独立产生若干正态随机变量，合在一起即得到多元标准正态随机向量。所以正态随机数/向量容易产生。

因为正态随机数容易产生，定理4（1）提供了利用多元标准正态随机向量从球面对称分布抽样的方法，定理4（2）提供了从球对称分布抽样的方法：

- 产生$\mathbf{x} \sim N_n(0, I_n)$；
- （球面均匀）令 $\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$，则$\mathbf{u} \sim U(S^{n-1})$;
- （球对称）为了产生 $\mathbf{y} \sim$球对称分布$f(\mathbf{y}) = h(\|\mathbf{y}\|)$的随机向量，产生$r \sim p(r) = \frac{2\pi^{n/2}}{\Gamma(n/2)} h(r) r^{n-1}$（假设容易产生），令$\mathbf{y} = \mathbf{u}r$。

第一讲命题1是定理2 的特殊情况：

$U(B^n)$随机表示的形式化表达：

$U(B^n) = U(S^{n-1}) \times [U(0,1)]^{1/n}$

例4. 假设 $\mathbf{x} \sim U(B^n)$, 即概率密度$f(\mathbf{x}) = \frac{1}{|B^n|} = \frac{n\Gamma(n/2)}{2\pi^{n/2}} 1_{(\|\mathbf{x}\| \le 1)}$。由定理2知

$$\frac{\mathbf{x}}{\|\mathbf{x}\|} \sim U(S^{n-1}), \frac{\mathbf{x}}{\|\mathbf{x}\|} \perp\!\!\!\perp \|\mathbf{x}\|，\ \|\mathbf{x}\| \sim p(r) = nr^{n-1} 1_{(0<r\le 1)}。$$

由定理3，若 $\mathbf{x} \sim U(B^n)$，则$\mathbf{x}$可表示为$\mathbf{x} \overset{d}{=} \mathbf{u}r$， $\mathbf{u} \sim U(S^{n-1})$,

$r \sim p(r) = nr^{n-1} 1_{(0<r\le 1)}$， 即 $r = U^{1/n}, U \sim U(0,1)$。

这可用于产生球均匀随机数：

球面均匀分布$U(S^{n-1})$和球均匀分布$U(B^n)$抽样

- 产生$\mathbf{x} \sim N_n(0, I_n)$，则$\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|} \sim U(S^{n-1})$;
- 产生$U \sim U(0,1)$,令$r = U^{1/n}$ $(r \sim p(r) = nr^{n-1}, 0 < r < 1)$;
- 令 $\mathbf{y} = \mathbf{u}r$， 则$\mathbf{y} \sim U(B^n)$ 。

注：拒绝抽样是一种简单易行的最常用的抽样/随机数产生方法，但这种方法不适用于生成 $U(B^n)$ 高维随机向量。

$U(B^n)$的拒绝抽样

- 产生$[-1,1]^n$均匀随机数$\mathbf{x}\sim U([-1,1]^n)$,
- 若$\|\mathbf{x}\| < 1$, 则$\mathbf{x}\sim U(B^n)$，即为所求；否则，拒绝。

但当n较大时单位球体积趋于0，大部分被拒绝，
拒绝抽样效率低下。上页最后的算法是一个$U(B^n)$抽样的高效方法。

附录：拒绝抽样Rejection sampling（acceptance-rejection sampling）

均匀分布

定义. 假设集合$C \subset R^n$的体积$|C| < \infty$，如果对任何$C_0 \subset C$,

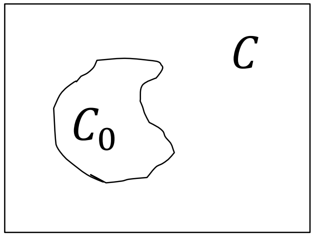
$$P(\mathbf{x} \in C_0) = |C_0| / |C|$$

则称 $\mathbf{x}$ 在C上均匀分布,记作$\mathbf{x} \sim U(C)$.

由均匀分布的定义，立得如下抽样基本原理

抽样基本原理1

假设 $\mathbf{x} \sim U(C)$，$C_0 \subset C$，则$\mathbf{x}|_{\mathbf{x} \in C_0} \sim U(C_0)$，即限制在子集$C_0$上也是均匀分布。



简单拒绝抽样

假设区域 C_0不规则，难以从中生成均匀随机数。为了产生 $\mathbf{x} \sim U(C_0)$，选取一个容易生成均匀随机数的规则区域C $\supset C_0$, 比如立方体， 产生$\mathbf{x} \sim U(C)$。若$\mathbf{x} \in C_0$，则$\mathbf{x}$即为所求；否则，舍弃/拒绝。

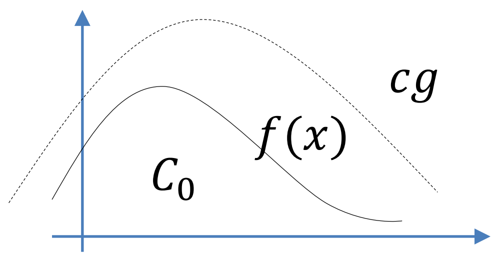
规则区域C包含C_0，C越小越好。上页规则区域 $C = [-1,1]^n \supset B^n$, 但C过大。

抽样基本原理2

考虑从概率密度f中抽样。若$(X,Y)\sim U(C_0)$，其中C_0为密度函数f的下方与横轴之间的区域（subgraph）

$$C_0 = sub(f) = \{(x,y): 0 \le y \le f(x)\}$$

则$X\sim f$。



拒绝抽样原理

为了从密度f抽样， 选取一个容易产生随机数的概率密度g, 假设存在$c \ge 1$, 使得 $cg \ge f$。 产生

$$(X,Y)\sim U(sub(cg)),$$

由基本原理2，若$Y \le f(X)$， 则X即为所求， 否则拒绝。

注意$(X,Y)\sim U\big(sub(cg)\big) \Leftrightarrow X\sim g, Y|X\sim U\big(0,cg(X)\big)$
$\Rightarrow U|X \triangleq \frac{Y}{cg(X)}|X\sim U(0,1)$，$U \perp\!\!\!\perp X$, 所以前述接受随机数的条件 “$Y \le f(X)$” $\Leftrightarrow U \le \frac{f(X)}{cg(X)}$。基于这些观察，重述上述拒绝抽样过程如下（更容易，不必真的从$sub(cg)$抽样）：

拒绝抽样

为了从密度f抽样，选取一个容易产生随机数的概率密度g, 假设存在$c \geq 1$, 使得 $cg \geq f$。

Rejection sampling：产生$X \sim g, U \sim U(0,1), U \perp\!\!\!\perp X$，若$U \leq \frac{f(X)}{cg(X)}$，则$X$即为所求，否则拒绝。

该方法首先产生一个$r.v.\ X \sim g$，计算$p = \frac{f(X)}{cg(X)}$.
再独立地产生一个$U \sim U(0,1)$，若$U \leq p$, 则接受X（这相当于抛一个正面概率为p的硬币，抛出正面就接受X, 抛出反面，拒绝）。

所以拒绝抽样方法在产生X后采取了随机化方式（抛一个正面概率为p的硬币）来决定是否接受X。这看起来很神秘，而且简单数学计算可以证明这样得到的X确实服从分布f，其背后原理如上页所述，并不神秘。