High-Dimensional Measures and Geometry
Lecture Notes from Jan 19, 2010

taken by Bernhard Bodmann

0 Course Information

Text: Michel Ledoux, The Concentration of Measure Phenomenon, American Mathematical Society, Providence, Rhode Island, 2001.

Office: Tu 1-2pm, Th 11-12 or by appointment

Email: bgb@math.uh.edu

Grade: Based on preparation of class notes in LaTeX, rotating note-takers

1 Introduction

Informally, the phenomenon of measure concentration is that in a given "high-dimensional" probability space, any "well-behaved" function is "almost" constant. We will study how to make this statement quantitative and precise and what consequences can be derived from this measure concentration. To get a taste of this subject, we start with a simple example, concentration on high-dimensional spheres.

1.1 Measure concentration on high-dimensional spheres

Consider the *unit sphere* $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, containing each vector $x \in \mathbb{R}^n$ with *Euclidean norm* $\|x\| = (\sum_{i=1}^n x_i^2)^{1/2} = 1$. We also define the *canonical inner product* between two vectors $x, y \in \mathbb{R}^n$ by $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$. Define the *metric distance* between $x, y \in \mathbb{S}^{n-1}$ by

$$d(x,y) = \arccos\langle x, y\rangle \in [0,\pi]\,.$$

1.1.1 Question. Why is the function $d : \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \to \mathbb{R}^+$ a metric?

Symmetry and positive definiteness are elementary. The distance $d(x,y)$ is the length of the shortest geodesic (great arc) between x and y. A direct proof of the triangle inequality can be found in M. Berger, Geometry II, Springer, N.Y., 1996. Otherwise, we can appeal to the fact that the shortest geodesic between two points minimizes length among all piecewise differentiable paths, and thus if the path had to pass through another point in between, then the length can only increase.

We also define the *rotation-invariant probability measure* μ on $\mathbb{S}^{n-1}$. This is induced on the sphere by the (left) Haar measure on the Lie group of all orientation-preserving rotations. For a given continuous function $f : \mathbb{S}^{n-1} \to \mathbb{R}$, we define the *median* m_f with respect to μ by the two properties

$$\mu(\{x \in \mathbb{S}^{n-1} : f(x) \geq m_f\} \geq \frac{1}{2}$$

and

$$\mu(\{x \in \mathbb{S}^{n-1} : f(x) \leq m_f\} \geq \frac{1}{2}\,.$$

1.1.2 Question. Why is this median well defined?

For each $\alpha \in \mathbb{R}$, define $L_\alpha = \{x \in \mathbb{S}^{n-1} : f(x) < \alpha\}$ and $U_\alpha = \{x \in \mathbb{S}^{n-1} : f(x) \geq \alpha\}$, then $g(\alpha) = \mu(L_\alpha)$ and $h(\alpha) = \mu(U_\alpha)$ are by definition monotonic in α and $g + h = 1$. Let the median m_f be given by $m_f = \inf\{\alpha : g(\alpha) \geq 1/2\}$, then it satisfies the second of the two claimed inequalities by the regularity of μ. Moreover, we also have $m_f = \sup\{\alpha : h(\alpha) \geq 1/2\}$, otherwise we would get a contradiction to $g+h = 1$. Using the regularity of μ, the first inequality follows.

1.1.3 Definition. *A function $f : \mathbb{S}^{n-1} \to \mathbb{R}$ is M-Lipshitz if*

$$|f(x) - f(y) \leq Md(x, y)$$

for all $x, y \in \mathbb{S}^{n-1}$.

We can now state a first result in the context concentration of measure.

1.1.4 Proposition. *Let $\mathbb{S}^{n-1}$ be the unit sphere in $\mathbb{R}^n$ and μ the rotation-invariant probability measure on $\mathbb{S}^{n-1}$. If $f : \mathbb{S}^{n-1} \to \mathbb{R}$ is 1-Lipshitz and m_f is its median with respect to μ, then for $\epsilon \geq 0$,*

$$\mu(\{x \in \mathbb{S}^{n-1} : |f(x) - m_f| \leq \epsilon\}) \geq 1 - \sqrt{\frac{\pi}{2}} e^{-\epsilon^2 (n-2)/2}\,.$$

Proof. Later. □

Note, the inequality gives no infor for "small" ϵ when n is held fixed. However, if we choose a sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ and let $\epsilon_n \to 0$ such that $\epsilon_n^2 n \to \infty$, then the measure $\mu(\{x \in \mathbb{S}^{n-1} : |f(x) - m_f| \leq \epsilon\}) \to 1$.

We will consider an example for a function f which is natural for the metric.

1.1.5 Question. Why is the function the function $f(x) = \langle x, a\rangle$, for $a \in \mathbb{S}^{n-1}$, 1-Lipshitz?

We have

$$|f(x) - f(y)| = |\langle x - y, a\rangle| \leq \|x - y\|\|a\| = \|x - y\| \leq d(x, y)\,,$$

because $\|x - y\|$ is the chordal distance, the length of the geodesic in $\mathbb{R}^n$ between x and y, whereas $d(x, y)$ measures the length of the shortest geodesic in the subset $\mathbb{S}^{n-1} \subset \mathbb{R}^n$.

1.1.6 Exercise. Derive a concentration inequality of the form stated in 1.1.4 in the special case $f(x) = \langle x, a\rangle$ for any $a \in \mathbb{S}^{n-1}$.

Without loss of generality, we can choose $a = (1, 0, 0, \dots, 0)$ after a suitable rotation of the sphere. The median of f is then identified as $m_f = 0$, because the "northern" hemisphere

$$\{x \in \mathbb{S}^{n-1} : \langle x, a\rangle \geq 0\} = \{x \in \mathbb{S}^{n-1} : x_1 \geq 0\}$$

has measure $1/2$ and so does its reflection about the origin.

Consider the image measure ρ_n on $[-1, 1]$ induced by projection μ with f. It is absolutely continuous with respect to the Lebesgue measure and has a Radon-Nikodym derivative

$$d\rho_n(t) = A_n(1 - t^2)^{(n-3)/2} dt\,,$$

with the normalization constant $A_n = \frac{\Gamma(n/2)}{\sqrt{\pi}\Gamma((n-1)/2)}$. We compute

$$\mu(\{x \in \mathbb{S}^{n-1} : |\langle x, a\rangle| \leq \epsilon\}) = \rho_n([-\epsilon, \epsilon])$$

use normalization and symmetry

$$\rho_n([-\epsilon, \epsilon]) = \int_{-\epsilon}^{\epsilon} A_n(1 - t^2)^{(n-3)/2} dt = 1 - 2\int_{\epsilon}^{1} A_n(1 - t^2)^{(n-3)/2} dt\,.$$

The integrand is estimated on the interval $[\epsilon, 1]$ by

$$1 - t^2 \leq e^{-t^2} \leq e^{-(n-3)\epsilon^2/2} e^{-(n-3)(t-\epsilon)}$$

where the exponent has been linearized about $t = \epsilon$. Further, extending the domain to $[\epsilon, \infty)$ (if $n > 3$) and performing the integration gives

$$\rho_n([-\epsilon, \epsilon]) \geq 1 - \frac{2A_n}{n-3} e^{-(n-3)\epsilon^2/2}\,.$$

This bound is of the same exponential type as in 1.1.4. Now the constant $2A_n/(n-3)$ can be estimated with the help of Stirling's approximation,

$$\frac{2A_n}{n-3} \approx \sqrt{\frac{2e}{n\pi}}$$

which vanishes as $n \to \infty$. Thus, this estimate is asymptotically slightly better than the result for the general case in 1.1.4.

We summarize this exercise informally by saying that the surface measure of the unit sphere in $\mathbb{R}^n$ is concentrated near the equator.

As we can see from the exercise, integration on the sphere can become quite involved. Fortunately, we can derive many concentration results with a detour through Gaussian measures in $\mathbb{R}^n$.

We will see that the Gaussian measure is in high dimensions close to the surface measure of an appropriately scaled sphere.

2 Wrapping the sphere in a Gaussian measure

The standard Gaussian measure γ_1 on $\mathbb{R}^1$ has the density $\frac{1}{\sqrt{2\pi}}e^{-t^2/2}$. It can be verified that this is a probability measure, and so is the standard Gaussian measure γ_n on $\mathbb{R}^n$ with density

$$\frac{1}{(2\pi)^{n/2}}e^{-\|x\|/2},$$

so for a measurable set $A \subset \mathbb{R}^n$,

$$\gamma_n(A) = \int_A \frac{1}{(2\pi)^{n/2}}e^{-\|x\|/2}dx\,.$$

We will find that the function $x \mapsto \|x\|$ is almost constant with respect to the Gaussian measure, with value $\|x\| \approx \sqrt{n}$. This means, with respect to the standard Gaussian measure, most vectors are close to the sphere of radius $\sqrt{n}$.