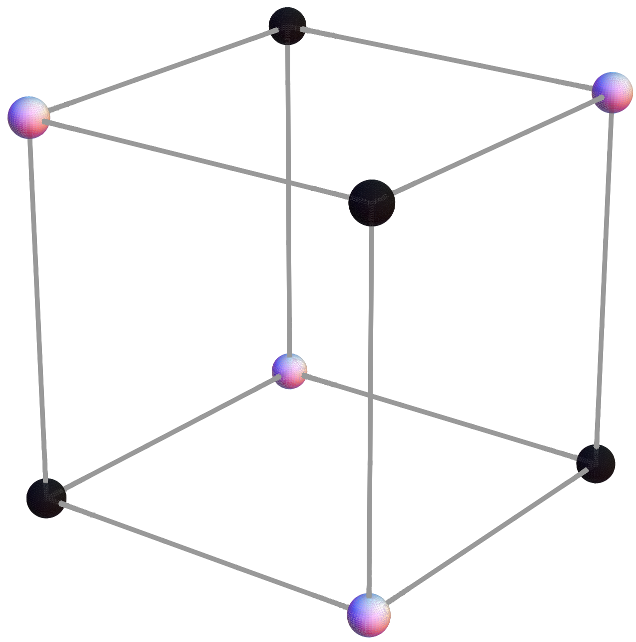
Noise sensitivity and Gaussian surface area

Keith Ball



The cube (in this talk) is $Q = \{-1, 1\}^n$ equipped with normalised counting measure.

A **Boolean** function f on Q is a function taking the values 1 and -1.

The **Noise sensitivity** of f measures how likely it is that the value of f will switch if we move our position in the cube a small amount.

So you pick a random corner X and then switch a randomly chosen εn of its coordinates to get a new point Y, and ask what is the probability that

$$f(X) \neq f(Y)?$$

The picture shows the “most” noise sensitive function: if you move one step you always change the value of f.

This function is a character on the group Q: the highest order character $X \mapsto \prod X_i$.

The least noise sensitive function is the constant function: the principal character. It makes more sense to look at functions with

$$\mathsf{P}(f(X) = 1) = \mathsf{P}(f(X) = -1) = 1/2.$$

Functions that put more weight on higher order characters, tend to be more noise sensitive.

Noise sensitivity is closely related to the study of the **influences** of variables on Boolean functions: the influence of the i^{th} variable is the chance that flipping this variable will change f.

So the sensitivity with $\varepsilon = \frac{1}{n}$ is the average influence.

The famous result of Kahn, Kalai and Linial states that for 50:50 functions, there must be a variable with influence at least

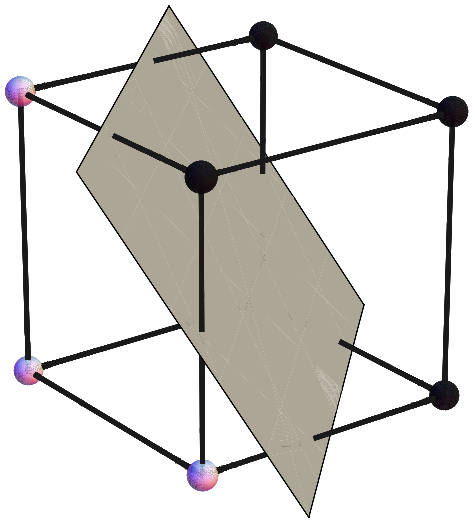
$$\frac{\log n}{n}$$

even though the average influence can be $1/n$.

Friedgut and Bourgain showed that if only few variables influence f then f is approximately a low order polynomial.

Talagrand estimates from below the expectation of the square root of the number of directions that flip f: so the reason that the average influence cannot be too small is not just a few bad points.

As for $\varepsilon = 1/n$, lower bounds on noise sensitivity don't tell us much: the focus on noise sensitivity is on upper estimates for functions of specific types: for example the sign of a linear function.



If you cut with the function $X \mapsto X_1$ then you have ε chance that your noisy coordinates include the first: so the sensitivity is ε.

The conjecture is that the worst direction is the main diagonal. The coordinates you switch have a reasonable chance of helping you by $\sqrt{\varepsilon n}$ so your point needs to be this close to having the same number of $+$ and $-$ coordinates. The chance of this is about $\sqrt{\varepsilon}$.

Peres proved a bound of order $\sqrt{\varepsilon}$. The sharp constant remains open.

If f is the indicator of the intersection of k half-spaces the sensitivity is at most $\sqrt{\varepsilon} k$ but the conjecture is the "usual" one: $\sqrt{\varepsilon}\sqrt{\log k}$.

A useful model for this problem is that of Gaussian noise sensitivity or Gaussian surface area.

If f is a Boolean function on $\mathbf{R}^n$ and X and Y are IID standard Gaussians then the $GNS(\varepsilon)$ is

$$\mathsf{P}(f(X) \neq f(\sqrt{1-\varepsilon}X + \sqrt{\varepsilon}Y)).$$

This is closely related to the Gaussian surface area of the set C where $f = 1$:

$$\int_{\partial C} g$$

where g is the standard Gaussian density.

If C has a smooth enough boundary the Gaussian surface area is

$$\lim_{\varepsilon \to 0} \frac{GNS(\varepsilon)}{\sqrt{\varepsilon}}.$$

So for the indicators of half-spaces we have the right dependence, $\sqrt{\varepsilon}$, for GNS.

Klivans, O'Donnell and Servedio use estimates for Gaussian surface area to measure algorithms for learning sets of different types and made a conjecture recently settled by D. Kane.

Theorem 1 (Ball). *If C is convex then its GSA is at most $4n^{1/4}$.*

Theorem 2 (Nazarov). *If C is the intersection of k half-spaces then its GSA is at most $\sqrt{\log k}$.*

Theorem 3 (Kane). *The GSA of ellipsoids is uniformly bounded.*

Nazarov also showed that the $n^{1/4}$ bound is sharp apart from the constant: for random sets with $\exp(\sqrt{n})$ facets.

To begin with, let's check the GSA of Euclidean balls.

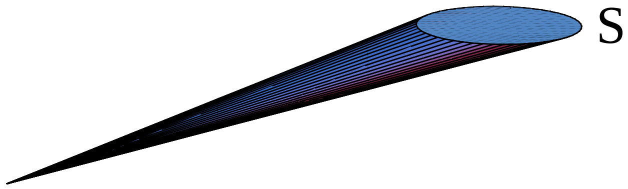
The ball of radius r has GSA

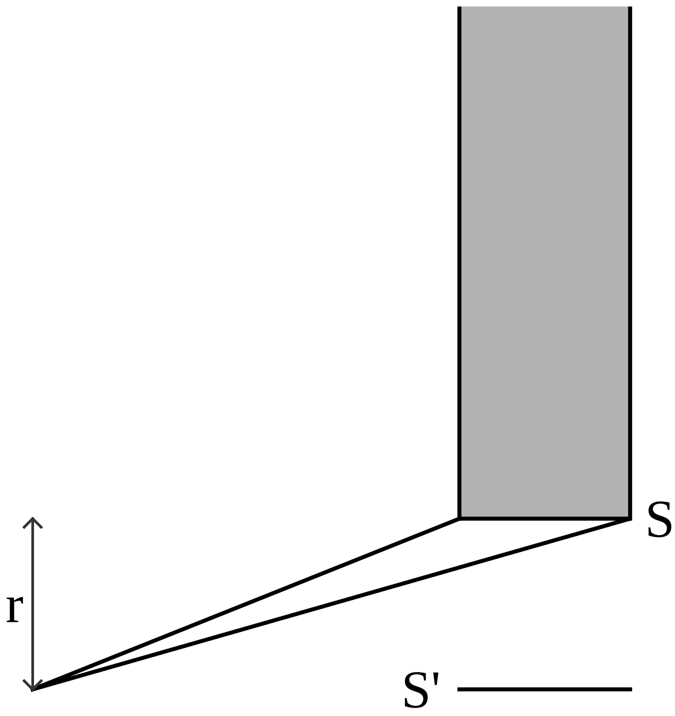
$$\frac{nr^{n-1}\pi^{n/2}}{\Gamma(n/2+1)}\,\frac{e^{-r^2/2}}{(\sqrt{2\pi})^n}$$

whose maximum occurs at $r = \sqrt{n-1}$ where the value is about $\frac{1}{\sqrt{\pi}}$.

It is also easy to check that the GSA for the (correctly sized) cube is $\sqrt{\log n}$ so Nazarov's estimate is sharp for this value of k as well.

Assume that 0 is inside C and consider a piece S of the surface of C.





The Gaussian volume of the shaded cylinder sitting above the surface is the product of the $(n-1)$-dimensional Gaussian volume of S' and the 1-dimensional Gaussian measure of the half-infinite interval.

This is

$$GSA(S)e^{r^2/2}\int_r^\infty e^{-x^2/2}\,dx \geq GSA(S)\frac{1}{1+r}.$$

Integrating over the surface of C we get

$$1-\gamma(C) \geq \int_{\partial C}\frac{1}{1+r(y)}g(y)$$

where $r(y)$ is the distance of the tangent plane at y, from 0.

From this we get an estimate when C is bounded by k hyperplanes.

$$1-\gamma(C) \geq \int_{\partial C} \frac{1}{1+r(y)} g(y)$$

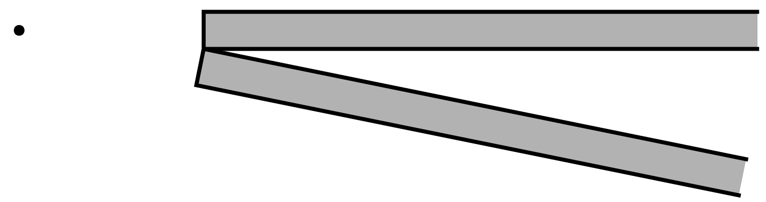
Hyperplanes at distance more than $\sqrt{2\log k}$ from the origin have Gaussian area at most $1/k$ and there are at most k of them.

For all points y on facets at distance less than $\sqrt{2\log k}$,

$$\frac{1}{1+r(y)} \geq \frac{1}{\sqrt{2\log k}}$$

so these contribute a GSA of at most $\sqrt{2\log k}$.

If we look at Nazarov's argument applied to a ball of radius about $\sqrt{n}$ (or a polytopal approximation to it) we have $r(y) = \sqrt{n}$ for all y on the surface and so it looks as though we will get GSA roughly $\sqrt{n}$.



The gaps don't get small.

To get an estimate of $n^{1/4}$ we use an argument motivated by Cauchy's integral formula for surface area

$$|\partial C| = c_n \int_{S^{n-1}} |P_\theta C| \, d\sigma.$$

Each projection is covered twice by the surface.

We try to find a measure μ on $\mathbf{R}^{n-1}$ so that for each small piece of surface S

$$GSA(S) = \int_{S^{n-1}} \mu(P_\theta C) \, d\sigma.$$

Then $GSA(C) \leq 2\mu(\mathbf{R}^{n-1})$.

The measure should have density $F(x) = f(|x|)$ and then for a small piece of surface centred at $r\phi$ with unit normal ψ the identity we want is

$$\frac{1}{\sqrt{2\pi}} e^{-r^2/2} = \int_{S^{n-1}} f\left(r\sqrt{1 - \langle \theta, \phi \rangle^2}\right) |\langle \theta, \psi \rangle| \, d\sigma.$$

This can't be true because of the the two angles ϕ and ψ. But we only need an inequality.

As long as f decreases on $[0, \infty)$ the right side is minimised when ϕ and ψ are orthogonal and in this case we get

$$\frac{2}{\pi} \int_0^{\pi/2} f(r \sin \theta) \sin^{n-1} \theta \, d\theta.$$

So we want

$$\frac{1}{\sqrt{2\pi}} e^{-r^2/2} = \frac{2}{\pi} \int_0^{\pi/2} f(r \sin \theta) \sin^{n-1} \theta \, d\theta.$$

$$\tilde{g}(r) = \frac{2}{\pi}\int_0^{\pi/2} f(r\sin\theta)\sin^{n-1}\theta\, d\theta.$$

The operator

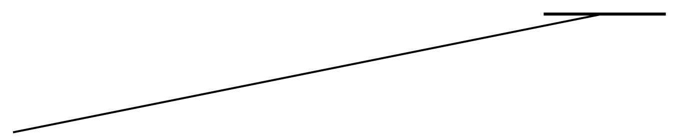
$$f \mapsto \frac{2}{\pi}\int_0^{\pi/2} f(.\sin\theta)\sin^{n-1}\theta\, d\theta$$

has polynomials as eigenfunctions so we can invert in a simple way.

$$\int_0^t f(r) r^{n-2}\, dr = t^{n-1}\int_0^{\pi/2} \tilde{g}(t\sin\theta)\sin^{n-2}\theta\, d\theta.$$

Now analyse f.

Nazarov showed that $n^{1/4}$ is sharp. The preceding argument shows that if we want near equality, the pieces of surface should have normal vectors almost perpendicular to their radius vectors.



The Gaussian measure lies at radius $\sqrt{n}$ so we want most of the surface to be at this distance from 0.

Nazarov's other argument shows that the bounding hyperplanes should be $n^{1/4}$ from 0.

These conditions are compatible.

Kane's argument estimates the noise sensitivity of an ellipsoid (or a solid whose surface is given by a polynomial of degree at most d).

f is the sign of a polynomial of degree d. X and Y are IID standard Gaussians and we want

$$p = \mathsf{P}\left(f(X) \neq f(\cos\theta\, X + \sin\theta\, Y)\right)$$

(where $\cos\theta = \sqrt{1-\varepsilon}$).

This is the same as

$$\mathsf{P}\left(f(\cos\theta\, X + \sin\theta\, Y) \neq f(\cos 2\theta\, X + \sin 2\theta\, Y)\right)$$

and

$$\mathsf{P}\left(f(\cos 2\theta\, X + \sin 2\theta\, Y) \neq f(\cos 3\theta\, X + \sin 3\theta\, Y)\right)$$

and so on.

So

$$np = \mathsf{E}\left(\mathbf{1}_{(f(Z_0)\neq f(Z_\theta))} + \cdots + \mathbf{1}_{(f(Z_{(n-1)\theta})\neq f(Z_{n\theta}))}\right).$$

The latter is at most the expectation of the number of sign changes of $f(Z_\phi)$ on the interval $[0, n\theta]$.

In the limit as $n \to \infty$ we get that p is at most

$\frac{\theta}{2\pi}$E(Number of sign changes of $f(Z_\phi)$ on $[0, 2\pi]$.

For each ω in the probability space

$$Z_\phi = \cos\phi\, X(\omega) + \sin\phi\, Y(\omega)$$

traces an ellipse as ϕ runs over $[0, 2\pi]$. We want to control the number of times this ellipse crosses the zero set of f.