Sergey Lototsky, USC

(Some) Gaussian Inequalities

Isoperimetric. The starting point is the `Brunn-Minkowski` inequality: if $\boldsymbol{\ell}$ is the Lebesgue measure on $\mathbb{R}^n$ and $A+B=\{a+b: a\in A, b\in B\}$ is the `Minkowski sum` of two sets, then, for (convex[1], later measurable[2]) sets A, B in $\mathbb{R}^n$ with non-empty interior,

$$\Big(\boldsymbol{\ell}(A+B)\Big)^{1/n} \geq \Big(\boldsymbol{\ell}(A)\Big)^{1/n} + \Big(\boldsymbol{\ell}(B)\Big)^{1/n}; \tag{1.1}$$

equality in (1.1) holds if and only if A and B are *homothetic*, that is, the same up to translation and dilation. When combined with the equality for the Lebesgue measure of the boundary ∂A of the set A,

$$\boldsymbol{\ell}(\partial A) = \liminf_{\varepsilon\to 0} \frac{\boldsymbol{\ell}(A+\varepsilon B_1) - \boldsymbol{\ell}(A)}{\varepsilon},\ B_1 = \{x=(x_1,\ldots,x_n)\in\mathbb{R}^n : x_1^2+\ldots+x_n^2<1\}, \tag{1.2}$$

inequality (1.1) leads to the classical `isoperimetric inequality`: if $\boldsymbol{\ell}(A)=\boldsymbol{\ell}(B_1)$, then $\boldsymbol{\ell}(\partial A)\geq \boldsymbol{\ell}(\partial B_1)$, *with equality if and only if* A is a ball of radius one; recall that $\boldsymbol{\ell}(\partial B_1) = 2\pi^{n/2}/\Gamma(n/2)$ and $\boldsymbol{\ell}(B_1)=\boldsymbol{\ell}(\partial B_1)/n$.

If $\boldsymbol{\mu}$ is the standard Gaussian measure on $\mathbb{R}^n$, then, similar to (1.2), define

$$\boldsymbol{\mu}^+(A) = \liminf_{\varepsilon\to 0} \frac{\boldsymbol{\mu}(A+\varepsilon B_1) - \boldsymbol{\mu}(A)}{\varepsilon} \tag{1.3}$$

and also the following functions

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},\ \ \Phi(x) = \int_{-\infty}^{x} \phi(t)\,dt,\quad I(y) = \phi\big(\Phi^{-1}(y)\big), \tag{1.4}$$

The result is a `Gaussian isoperimetric inequality`[3]

$$\boldsymbol{\mu}^+(A) \geq I\big(\boldsymbol{\mu}(A)\big), \tag{1.5}$$

with equality if and only if A is a half-space $\{x\in\mathbb{R}^n : x_1a_1+\ldots+x_na_n\leq r\}$ for some fixed $a\in\mathbb{R}^n$ and $r\in\mathbb{R}$. Two equivalent forms of (1.5) are

$$\Phi^{-1}\big(\boldsymbol{\mu}(A+\varepsilon B_1)\big) \geq \Phi^{-1}\big(\boldsymbol{\mu}(A)\big) + \varepsilon,\ \varepsilon>0 \tag{1.6}$$

and[4]

$$I\big(\mathbb{E}(f(Z))\big) \leq \mathbb{E}\sqrt{I^2(f(Z)) + |\nabla f(Z)|^2}; \tag{1.7}$$

in (1.7), I is the function from (1.4), Z is a standard Gaussian random vector in $\mathbb{R}^n$, and f is a continuously differentiable function satisfying $0<f(x)<1$.

The main corollaries of (1.5)–(1.7) are

- `the (Gaussian) log-Sobolev inequality`[5]

$$\mathbb{E}\left(|g(Z)|^2 \ln |g(Z)|\right) \leq \mathbb{E}|\nabla g(Z)|^2 \tag{1.8}$$

  for a continuously differentiable function g satisfying $\mathbb{E}g^2(Z)=1$ [can be derived from (1.7)];
- various `concentration inequalities`, such as

$$\mathbb{P}\big(f(Z) > M_f + t\big) \leq 1 - \Phi(t/\sigma),\ t>0, \tag{1.9}$$

  where $|f(x)-f(y)|\leq \sigma|x-y|,\ x,y\in\mathbb{R}^n$ and M_f is the median of the random variable $f(Z)$ [this was part of Borell's paper from 1975].

---

[1] Brunn (1887, $n=3$), Minkowski (1896)
[2] Lyusternik (1935)
[3] Borell (1975), Tsyrelson-Sudakov (1974)
[4] Bobkov (1997)
[5] Gross (1975)

Because (1.5) does not involve n, an extension to a locally convex topological space is (almost) immediate, *as long as the init ball* B_1 *in* (1.3) *is in the Cameron-Martin space* H_μ of $\boldsymbol{\mu}$.

Supremum of a Gaussian process. We have a real-valued, zero-mean Gaussian process $X = X(t)$, $t \in \mathbf{T}$, indexed by an arbitrary set $\mathbf{T}$ with at least two elements. The key objects turn out to be

$$\sigma_X^2 = \sup_{t\in\mathbf{T}} \mathbb{E}|X(t)|^2, \quad \rho_X(s,t) = \sqrt{\mathbb{E}|X(t)-X(s)|^2}. \tag{1.10}$$

Note that $\rho_X(t,t) = 0$, but it is possible to have $\rho_X(s,t) = 0$ for some $t \neq s$. Still, by Minkowski's inequality (for $p = 2$), the function ρ satisfies the triangle inequality and thus defines a (pseudo) distance/metric on $\mathbf{T}$ (sometimes called the *canonical distance/metric*).

To simplify notations, write

$$X^*_{\mathbf{T}} = \sup_{t\in\mathbf{T}} X(t). \tag{1.11}$$

Some results, such as Slepian's inequality (see below), do not hold if $X^*_{\mathbf{T}}$ is replaced with $\sup_{t\in\mathbf{T}} |X(t)|$. For many other purposes, the simple relations

$$\mathbb{P}(X^*_{\mathbf{T}} > r) \le \mathbb{P}(\sup_{t\in\mathbf{T}} |X(t)| > r) \le 2\mathbb{P}(X^*_{\mathbf{T}} > r), \;\; r \in \mathbb{R}, \tag{1.12}$$

$$\mathbb{E}X^*_{\mathbf{T}} \le \mathbb{E}\sup_{t\in\mathbf{T}} |X(t)| \le 2\Big(\mathbb{E}X^*_{\mathbf{T}} + \inf_{t\in\mathbf{T}} \sqrt{\mathbb{E}|X(t)|^2}\Big), \tag{1.13}$$

are enough to go from $X^*_{\mathbf{T}}$ to $\sup_{t\in\mathbf{T}} |X(t)|$. The reason for the second inequality in (1.12) is that, for a zero-mean Gaussian process $X = X(t)$, the process $-X$ has the same covariance function and hence the same distribution. By noticing that if $X^*_{\mathbf{T}} < 0$, then $X^*_{\mathbf{T}} = -\inf_{t\in\mathbf{T}} |X(t)|$, we can combine (1.12) with the equality

$$\mathbb{E}Y = \int_0^{+\infty} \mathbb{P}(Y > r)\, dr - \int_{-\infty}^0 \mathbb{P}(Y \le r)\, dr$$

to get (1.13).

According to L. Shepp (et al.)[6], $\mathbb{P}(X^*_{\mathbf{T}} < \infty)$ is either 0 or 1; if $\mathbb{P}(X^*_{\mathbf{T}} < \infty) = 1$, then

$$\lim_{u\to+\infty} \frac{\ln \mathbb{P}(X^*_{\mathbf{T}} > u)}{u^2} = -\frac{1}{2\sigma_X^2}. \tag{1.14}$$

In particular,

$$\mathbb{P}(X^*_{\mathbf{T}} < \infty) = 1 \;\; \Rightarrow \;\; \mathbb{E}e^{\varepsilon(X^*_{\mathbf{T}})^2} < \infty \;\text{ for all sufficiently small }\; \varepsilon > 0.$$

One collection of results starts with the `Dudley integral`

$$\mathcal{D}(r) = \int_0^r \sqrt{\ln N(r)}\, dr, \tag{1.15}$$

where $N(r)$ is the minimal number of open ball of radius r required to cover $\mathbf{T}$; the radius of the balls is computed with respect to ρ_X. Then[7]

$$\mathbb{E}X^*_{\mathbf{T}} \le 4\sqrt{2}\,\mathcal{D}(\sigma_X/2); \tag{1.16}$$

if the right-hand side of (1.16) is finite, then X has a continuous modification on the (pseudo) metric space $(\mathbf{T}, \rho_X)$; if the process X is stationary and $\mathbf{T} = [a,b] \subset \mathbb{R}$, then continuity of X implies that $\mathcal{D}$ is finite. Note that, for (1.15) to be finite, it is necessary (but not sufficient) to have $\mathbf{T}$ compact with respect to ρ_X.

Inequality (1.16) extends to any zero-mean process $X = X(t)$ that is *sub-Gaussian* with respect to the canonical [or some other...] metric ρ_X: $\mathbb{E}e^{\lambda(X(t)-X(s))} \le e^{\lambda^2\rho_X^2(t,s)/2}$. In other words, instead of a Gaussian process indexed by an arbitrary set $\mathbf{T}$, the story can begin with a sub-Gaussian process indexed by a (pseudo) metric space $(\mathbf{T}, \rho_X)$.

[6] For example, Landau-Shepp (1970)

[7] Dudley (1967)

The `Borell-TIS`[8] inequality is somewhat different from (1.16) and is more along the lines of (1.9): assuming that $\mathbb{P}(X^*_{\mathbf{T}} < \infty) = 1$ and using the notation $m^* = \mathbb{E}X^*_{\mathbf{T}}$,

$$\mathbb{P}(X^*_{\mathbf{T}} - m^* > r) \le e^{-r^2/(2\sigma_X^2)},\ r > 0. \tag{1.17}$$

As a consequence, we can get a more detailed information about continuity of X on $(\mathbf{T}, \rho_X)$: the sample paths of X are continuous if and only if if $\lim_{\varepsilon\to 0} \psi(\epsilon) = 0$, where

$$\psi(\varepsilon) = \mathbb{E} \sup_{\rho_X(t,s)<\varepsilon} (X(t) - X(s)). \tag{1.18}$$

If (1.18) holds then, with probability one and for all sufficiently small $\delta > 0$,

$$\sup_{\rho_X(t,s)<\delta} |X(t) - X(s)| \le \psi(\delta)|\ln \psi(\delta)|^p, \quad p > 0.$$

Comparison inequalities deal with two zero-mean Gaussian processes X and Y. Using the notations (1.10) and (1.11), the two main results are as follows:

- **Slepian's inequality.**[9] If $\mathbb{E}|X(t)|^2 = \mathbb{E}|Y(t)|^2$ and $\rho_X(s,t) \le \rho_Y(s,t)$, $s,t \in \mathbf{T}$, then $Y^*_{\mathbf{T}}$ *stochastically dominates* $X^*_{\mathbf{T}}$:

$$\mathbb{P}(X^*_{\mathbf{T}} > r) \le \mathbb{P}(Y^*_{\mathbf{T}} > r),\ r \in \mathbb{R}. \tag{1.19}$$

  In particular,

$$\mathbb{E}X^*_{\mathbf{T}} \le \mathbb{E}Y^*_{\mathbf{T}}. \tag{1.20}$$

- **Fernique-Sudakov inequality.**[10] If we only have $\rho_X(s,t) \le \rho_Y(s,t)$ (but not $\mathbb{E}|X(t)|^2 = \mathbb{E}|Y(t)|^2$), then we still have (1.20) (but not necessarily (1.19)).

An immediate consequence of Fernique-Sudakov is that $\mathbb{E}X^*_{\mathbf{T}} \ge 0$, and the inequality is strict unless $\rho_X(s,t) = 0$ for all $t, s \in \mathbf{T}$.

While Slepian's inequality is more informative, the condition $\mathbb{E}|X(t)|^2 = \mathbb{E}|Y(t)|^2$ makes the result hard to use, whereas Fernique-Sudakov readily applies to many familiar Gaussian processes. For example, if $X = W^H$ is the *fractional Browian motion* with Hurst parameter $H \in (0,1)$, then $\rho_X(t,s) = |t-s|^H$, and we immediately conclude that the function $H \mapsto \mathbb{E}\sup_{0<t<1} W^H(t)$ is decreasing in H.

Recall that a Gaussian measure $\boldsymbol{\mu}$ on a locally convex topological space $\mathbf{X}$ is called `centered` if, for every bounded linear functional f on $\mathbf{X}$, the (Gaussian) random variable $f(x)$ on $(\mathbf{X}, \mathcal{B}(\mathbf{X}), \boldsymbol{\mu})$ has zero mean.

Anderson's inequality for a centered Gaussian measure $\boldsymbol{\mu}$ on a locally convex topological space $\mathbf{X}$, a convex symmetric set $A \subset \mathbf{X}$ and a fixed element $x \in \mathbf{X}$,

$$\boldsymbol{\mu}(A + x) \le \boldsymbol{\mu}(A),$$

has an alternative, and more detailed version in $\mathbb{R}^n$: if $f = f(x)$ is a pdf such that $f(x) = f(-x)$ and the sets $\{x : f(x) \ge t\}$ are convex for every $t \ge 0$, then

$$\int_A f(x + cy)\, dx \ge \int_A f(x + y)\, dx$$

for every convex symmetric set A and every $c \in (0,1)$. For the lower bound on shifted Gaussian measure, there is an inequality due to Borell (1977):

$$\boldsymbol{\mu}(A + h) \ge e^{-\|h\|^2_{H_\mu}}\, \boldsymbol{\mu}(A), \quad h \in H_\mu,$$

where $\boldsymbol{\mu}$ is a centered Gaussian measure on a locally convex topological space $\mathbf{X}$, $A \subset \mathbf{X}$ is a measurable symmetric (but not necessarily convex) set, and H_μ is the Cameron-Martin space of $\boldsymbol{\mu}$.

[8]Borell (1975), Tsyrelson-Ibragimov-Sudakov (1975)
[9]Slepian (1963)
[10]Fernique (1974), Sudakov (1970)

Gaussian Correlation Inequality for a centered Gaussian measure $\boldsymbol{\mu}$ and symmetric convex sets A, B,

$$\boldsymbol{\mu}(A \cap B) \geq \boldsymbol{\mu}(A)\boldsymbol{\mu}(B),$$

used to be a conjecture for about 40 years, and was finally proved in 2014 by Thomas Royen.

An example of the **Gaussian Poincaré inequality** is

$$\mathbb{E}|f(Z)|^2 \leq \mathbb{E}|\nabla f(Z)|^2, \tag{1.21}$$

where Z is a standard Gaussian random variable on $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function satisfying $\mathbb{E}f(Z) = 0$.

An example of a **Gaussian hypercontractivity inequality** is

$$\left(\mathbb{E}|T_t f(Z)|^q\right)^{1/q} \leq \left(\mathbb{E}|f(Z)|^p\right)^{1/p}, \tag{1.22}$$

where $q > p > 1$, $e^{-t} < \sqrt{\frac{p-1}{q-1}}$, Z is a standard Gaussian random variable on $\mathbb{R}^n$ and, for a (bounded measurable) function $f : \mathbb{R}^n \to \mathbb{R}$ and $t > 0$,

$$T_t f(x) = \mathbb{E}f\left(X_x(t)\right), \quad X_x(t) \sim \mathcal{N}\big(xe^{-t}, (1 - e^{-2t})I_{n\times n}\big). \tag{1.23}$$

One can use (1.23) to prove (1.21); a suitable version of (1.22) can lead to (1.8).

Historical comments.

Theodore Wilbur Anderson (1918–2016) specialized in multivariate analysis and was professor at Columbia (1946-67) and Stanford (1967-88).

Sergey Bobkov was a student of Sudakov and is now professor at the University of Minnesota.

Christer Borell got his Ph.D. in 1974 from Uppsala university, with dissertation *Convex measures on infinite-dimensional spaces*; he is Professor Emeritus at Chalmers University of Technology in Göteborg, Sweden.

German mathematician Karl Hermann Brunn (1862–1939) worked in convex geometry and knot theory; he was born in Rome.

American mathematician Richard Mansfield Dudley (1938–2020) was Putnam Fellow and spent most of his career at MIT (1967–2015).

French mathematician Xavier Fernique (1934–2020) was the Ph.D. advisor of Michel Ledoux.

American mathematician Leonard Gross was born in 1931 and spent most of his career at Cornel (from 1960 on), where he supervised over 30 Ph.D. students.

Born in 1932, Ildar Abdulovich Ibragimov was an invited speaker at ICM in 1966 and is still active in math research as a senior member of the laboratory of statistical methods at the Steklov Institute in Saint Petersburg.

Mathematician Hermann Minkowski (1864–1909) was born in the southern suburbs of Kaunas (at that time, the Polish part of the Russian empire), studied in Königsberg (now Kaliningrad), and worked in Germany; he is famous for many important results, and so all the countries involved (Germany, Lithuania, Poland, Russia) occasionally claim him as their own.

Born in 1947, Thomas Royen worked as a statistician at the pharmaceutical company Hoechst AG (1977-85), and then taught mathematics and statistics at the University of Applied Sciences Bingen in Rhineland-Palatinate (1985-2010); apparently, he never used TEX to write his papers.

American mathematician Lawrence Alan Shepp (1936–2013) was a Ph.D. student of W. Feller at Princeton and had various positions at Bell Labs, Rutgers, UPenn (Warton), Stanford, and the Columbia Presbyterian Hospital (NYC). Two of his papers, including one in the *Annals of Statistics*, were published in 2017.

American mathematician David S. Slepian (1923–2007) worked at Bell Labs and made fundamental contributions to coding theory. His father, Joseph Slepian, got a Ph.D. from Harward under D. Birkhoff; his grandparents immigrated from Russia.

Born in 1934, Vladimir Nikolaevich Sudakov keeps a surprisingly low profile. Some on-line sources suggest that he is "academic brother" of I. A. Ibragimov (both were students of Yu. V. Linnik), and he was his colleague at the Steklov Institute in Saint Petersburg.

Boris Semyonovich Tsirelson (1950–2020) was a Ph.D. student of I. A. Ibragimov.