

1 Functions of normal random variables

It is of course impossible to even think the word Gaussian without immediately mentioning the most important property of Gaussian processes, that is concentration of measure.

M. Talagrand¹

1a	Gaussian isoperimetry and related inequalities .	1
1b	Application: decoding error probability	3
1c	Useful special cases	3
1d	Application: frozen disorder models	4

1a Gaussian isoperimetry and related inequalities

1a1 Definition. (a) The standard Gaussian measure γ^1 on \mathbb{R} , called also the standard normal distribution $N(0, 1)$, is the probability measure

$$\gamma^1(A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \quad \text{for Lebesgue measurable } A \subset \mathbb{R}.$$

(b) The standard Gaussian measure γ^n on \mathbb{R}^n , called also the standard multi-normal distribution, is the probability measure $\gamma^1 \times \cdots \times \gamma^1$, that is,

$$\gamma^n(A) = \int_A (2\pi)^{-n/2} e^{-|x|^2/2} dx \quad \text{for Lebesgue measurable } A \subset \mathbb{R}^n.$$

Here and below, $|\cdot|$ is the Euclidean norm.

A random vector $(X_1, \dots, X_n) = X : \Omega \rightarrow \mathbb{R}^n$ is distributed γ^n if and only if X_1, \dots, X_n are independent $N(0, 1)$ random variables.

Let μ be a nonatomic probability measure on \mathbb{R} , and ν an arbitrary probability measure on \mathbb{R} . Then there exists an increasing $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f[\mu] = \nu$, that is, f sends μ into ν . Such f is unique up to the values at points of discontinuity; it is unique if the support of ν is connected, which holds in all cases treated below.

In particular, for every random variable $\xi : \Omega \rightarrow \mathbb{R}$ there exists an increasing $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(\xi)$ is distributed like ξ if ξ is distributed $N(0, 1)$.

¹See page 189 of “Mean field model for spin glasses: a first course”, Lecture Notes in Math. **1816** (2003), 181–285.

1a2 Theorem. ¹ Let a function $\xi : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfy the Lipschitz condition with constant 1:

$$\begin{aligned} \xi \in \text{Lip}(1), \quad \text{that is,} \\ \forall x, y \in \mathbb{R}^n \quad |\xi(x) - \xi(y)| \leq |x - y|. \end{aligned}$$

Then $\xi[\gamma^n] = f[\gamma^1]$ for an increasing $f : \mathbb{R} \rightarrow \mathbb{R}$, $f \in \text{Lip}(1)$.

The same holds for $\text{Lip}(C)$ (by rescaling).

Note that both the supremum and the infimum of $\text{Lip}(1)$ functions is a $\text{Lip}(1)$ function (if finite).

1a3 Theorem. ² Let a function $\xi : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex:

$$\forall x, y \in \mathbb{R}^n \quad \forall c \in [0, 1] \quad \xi(cx + (1 - c)y) \leq c\xi(x) + (1 - c)\xi(y).$$

Then $\xi[\gamma^n] = f[\gamma^1]$ for a convex increasing $f : \mathbb{R} \rightarrow \mathbb{R}$.

If ξ is $\text{Lip}(C)$ and convex then f is (increasing and) $\text{Lip}(C)$ and convex. In particular, this holds for a (semi)norm:

$$\xi(x) = \sup_{y \in Y} |\langle x, y \rangle|$$

for a bounded $Y \subset \mathbb{R}^n$; in this case $C = \sup_{y \in Y} |y|$. In addition, $\forall x \in \mathbb{R} \quad f(x) \geq Cx$ (for a seminorm).

However, for a seminorm ξ it is natural to introduce an increasing $g : (0, \infty) \rightarrow (0, \infty)$ such that ξ is distributed like $g(|\zeta|)$, $\zeta \sim N(0, 1)$. It appears that

$$\text{the function } a \mapsto \frac{g(a)}{a} \text{ is decreasing on } (0, \infty),$$

which is called the S-inequality.³

Another result for a seminorm ξ may be formulated in terms of a random variable $\eta \sim \text{Exp}(1)$, that is, $\forall a \in [0, \infty) \quad \mathbb{P}(\eta \geq a) = e^{-a}$, and a *decreasing* function $h : (0, \infty) \rightarrow (0, \infty)$ such that $\xi \sim h(\eta)$. It appears that

$$\text{the function } \ln h(\cdot) \text{ is convex on } (0, \infty),$$

which is called the B-inequality.⁴

It is conjectured that for all seminorms ξ_1, ξ_2 ,

$$\mathbb{P}(\xi_1 \leq 1 \text{ and } \xi_2 \leq 1) \geq \mathbb{P}(\xi_1 \leq 1)\mathbb{P}(\xi_2 \leq 1).$$

This ‘‘correlation conjecture’’ is a 40-years old open problem!⁵

¹V. Sudakov, B. Tsirelson 1974, and independently C. Borell 1975.

²A. Ehrhard 1983.

³R. Latała, K. Oleszkiewicz 1999.

⁴D. Cordero-Erausquin, M. Fradelizi, B. Maurey 2004.

⁵See: G. Schechtman, Th. Schlumprecht, J. Zinn (1998), ‘‘On the Gaussian measure of the intersection’’, *Annals of Probability* **26**:1, 346–357.

1b Application: decoding error probability

A code of length n is a finite set $C \subset \mathbb{R}^n$; each $c \in C$ is a codevector. A sender submits a given codevector c_0 to the additive white Gaussian noise channel, and it gets corrupted to $c_0 + \sigma X$ where X is distributed γ^n and σ is the channel noise value. A receiver chooses the codevector c closest to the received vector $c_0 + \sigma X$ (which means maximum-likelihood decoding).

Consider the error probability

$$P_{\text{error}}(\sigma) = \mathbb{P}(c \neq c_0)$$

and the minimal distance

$$D = \min_{c \in C, c \neq c_0} |c - c_0|$$

(for given C and c_0).

1b1 Theorem. There exists an increasing Lip(1) convex function $f : \mathbb{R} \rightarrow (0, \infty)$ such that

$$\forall \sigma > 0 \quad P_{\text{error}}(\sigma) = \mathbb{P}\left(f(\zeta) \geq \frac{D}{2\sigma}\right)$$

where $\zeta \sim N(0, 1)$.

Defining the critical noise value σ_c by $P_{\text{error}}(\sigma_c) = 0.5$ we get¹

$$\forall \sigma \in (0, \sigma_c) \quad P_{\text{error}}(\sigma) \leq \mathbb{P}\left(\zeta \geq \frac{D}{2}\left(\frac{1}{\sigma} - \frac{1}{\sigma_c}\right)\right).$$

1c Useful special cases

Let $y_1, \dots, y_n \in \mathbb{R}^d$, $\forall k |y_k| \leq 1$.

The function

$$(1c1) \quad \xi(x) = \max_k \langle x, y_k \rangle$$

is Lip(1) and convex. The function

$$(1c2) \quad \xi(x) = \max_k |\langle x, y_k \rangle|$$

¹This is Theorem 1 of: J.-P. Tillich, G. Zemor (2004) "The Gaussian isoperimetric inequality and decoding error probabilities for the Gaussian channel", IEEE Transactions on Information Theory **50**:2, 328–331.

is Lip(1) and a seminorm.

For $p \in [1, \infty)$ the function

$$(1c3) \quad \xi(x) = \left(\frac{1}{n} \sum_k |\langle x, y_k \rangle|^p \right)^{1/p}$$

is Lip(1) and a seminorm.

For $c \in \mathbb{R}$, $c \neq 0$, the function

$$(1c4) \quad \xi(x) = \frac{1}{|c|} \ln \sum_k e^{c\langle x, y_k \rangle}$$

is Lip(1) and convex.

More generally, for a nonempty subset A of the unit ball of \mathbb{R}^d and a probability measure μ on the unit ball of \mathbb{R}^d we may generalize (1c1) to $\xi(x) = \sup_{y \in A} \langle x, y \rangle$, (1c2) to $\xi(x) = \sup_{y \in A} |\langle x, y \rangle|$, (1c3) to $\xi(x) = \left(\int |\langle x, y \rangle|^p \mu(dy) \right)^{1/p}$, and (1c4) to $\xi(x) = \frac{1}{|c|} \ln \int e^{c\langle x, y \rangle} \mu(dy)$.

For every nonempty $A \subset \mathbb{R}^d$ the function

$$(1c5) \quad \xi(x) = \inf_{y \in A} |x - y|$$

is Lip(1). If A is convex then the function is also convex.

Infinite dimension will be discussed later.

1d Application: frozen disorder models

Let $X_{k,l}$ for $0 \leq k \leq m$ and $-k \leq l \leq k$ be independent $N(0, 1)$ random variables. Every path $L = (l_0, \dots, l_m)$ such that $l_0 = 0$ and $l_{k+1} - l_k = \pm 1$ (for $k = 0, \dots, m-1$) leads to a random variable

$$X_L = \sum_{k=0}^m X_{k, l_k}.$$

FIRST-TIME PERCOLATION

The random variable

$$\xi_m = \frac{1}{\sqrt{m}} \max_L X_L$$

is a Lip(1) convex function of $(X_{k,l})_{k,l}$. Gaussian concentration helps to prove the strict inequality

$$\lim_{m \rightarrow \infty} \frac{1}{\sqrt{m}} \mathbb{E} \xi_m < \sqrt{2 \ln 2},$$

and the same for d -dimensional l , with $\sqrt{2 \ln(2d)}$ in the right-hand side.¹

¹See Theorem 1.2 in: Y. Hu, Q.-M. Shao (2009), "A note on directed polymers in Gaussian environments", *Electronic Communications in Probability* **14**, 518–528.

DIRECTED POLYMER

For $\beta > 0$ the random variable

$$Z_{m,\beta} = \frac{1}{2^m} \sum_L e^{\beta X_L}$$

is the so-called partition function of the directed polymer in Gaussian random environment; β is the inverse temperature. (The energy of L is, by definition, $-X_L$; the probability of L , given $(X_{k,l})_{k,l}$, is $\frac{1}{Z_{m,\beta} 2^m} e^{\beta X_L}$.) The limit of $\frac{1}{m} \mathbb{E} \ln Z_{m,\beta}$ as $m \rightarrow \infty$ is of interest;² in dimension 3 it is equal to $\beta^2/2$ for β below a critical value, but strictly less than $\beta^2/2$ for β above the critical value.

The random variable

$$\xi = \frac{1}{\beta \sqrt{m}} \ln Z_{m,\beta}$$

is a Lip(1) convex function of $(X_{k,l})_{k,l}$.

SPIN GLASS

Let $X_{k,l}$ for $1 \leq k < l \leq m$ be independent $N(0, 1)$ random variables. Every $(\sigma_1, \dots, \sigma_m) = \sigma \in \{-1, +1\}^m$ leads to a random variable

$$X_\sigma = \sum_{k < l} \sigma_k \sigma_l X_{k,l}.$$

For $\beta > 0$ the random variable

$$Z_{m,\beta} = \frac{1}{2^m} \sum_\sigma e^{\beta X_\sigma / \sqrt{m}}$$

is the partition function of the so-called Sherrington-Kirkpatrick model for spin glasses (“SK model”, nonlocal; the energy of σ is, by definition, $-X_\sigma / \sqrt{m}$; the probability of σ , given $(X_{k,l})_{k,l}$, is $\frac{1}{Z_{m,\beta} 2^m} e^{\beta X_\sigma / \sqrt{m}}$). The limit of $\frac{1}{m} \mathbb{E} \ln Z_{m,\beta}$ (as $m \rightarrow \infty$) is of interest;³ it is equal to $\beta^2/4$ for $\beta < 1$, but strictly less than $\beta^2/4$ for $\beta > 1$.⁴

The random variable

$$\xi = \frac{1}{\beta} \sqrt{\frac{2}{m-1}} \ln Z_{m,\beta}$$

²See Hu and Shao, Sect. 1.1.

³See Talagrand, page 190.

⁴Rather, $\limsup(\dots) < \beta^2/4$.

is a $\text{Lip}(1)$ convex function of $(X_{k,l})_{k,l}$. Gaussian concentration shows that $\frac{1}{m} \ln Z_{m,\beta}$ is usually close to its expectation. This random function of β “is convex and has small fluctuations. The quantity of information that can be extracted from this simple fact is amazing.” (Talagrand, page 191)

References

- [1] V.I. Bogachev, *Gaussian measures*, AMS 1998.
- [2] R. Latała, *On some inequalities for Gaussian measures*, Proc. Intern. Congress of Mathematicians (ICM 2002), vol. II, 813–822.