Large deviations of Gaussian vectors

Let X be a centered Gaussian random vector, on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with values in a real separable Banach space E equipped with its Borel σ-algebra $\mathcal{B}$, and with norm $\| \cdot \|$.

It is a consequence of the sharp integrability of the norms of Gaussian random vectors (cf. [1]) that

$$\lim_{t \to \infty} t^2 \log \mathbb{P}\big(\|X\| \geq t \big) \, = \, - \frac{1}{2\sigma^2} \tag{1}$$

where

$$\sigma^2 \, = \sup_{\xi \in E^*, \|\xi\| \leq 1} E\big(\langle \xi, X \rangle^2 \big). \tag{2}$$

This result is actually a particular case of a more general large deviation principle for the family of laws of εX as $\varepsilon \to 0$, providing further knowledge on tail behaviors.

The post briefly presents this large deviation theorem. General references on (Gaussian) large deviations include [14, 8, 7, 11, 6, 13] etc.

Table of contents

Drafted by M. L. **v1** December 2023

1 Rate function

Given a centered Gaussian random vector X on $(\Omega, \mathcal{A}, \mathbb{P})$ with values in E, its law μ on the Borel sets of E gives rise to an abstract Wiener space structure $(E, \mathcal{H}, \mu)$, in which the Hilbert space $\mathcal{H} \subset E$, with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, is the reproducing kernel Hilbert space associated to the covariance structure of μ (cf. [2]).

For the example of the Wiener measure μ on the Banach space $E = C([0,1])$ of real continuous functions on $[0,1]$, law of a standard Brownian motion or Wiener process $W = (W(t))_{t \in [0,1]}$, the reproducing kernel Hilbert space $\mathcal{H}$ is identified as the subspace of $E = C([0,1])$ consisting of the absolutely continuous functions $h : [0,1] \to \mathbb{R}$, with almost everywhere derivative h' in $\mathrm{L}^2([0,1])$ (for the Lebesgue measure), and with

$$|h|_{\mathcal{H}} = \left(\int_0^1 h'(t)^2 \, dt \right)^{1/2}.$$

The rate function $\mathcal{I} : E \to [0, +\infty]$ which will govern the large deviation properties of εX as $\varepsilon \to 0$ is defined as

$$\mathcal{I}(x) = \begin{cases} \frac{1}{2}|x|_{\mathcal{H}}^2 & \text{if} \quad x \in \mathcal{H}, \\ +\infty & \text{if} \quad x \notin \mathcal{H}. \end{cases} \tag{3}$$

In the large deviation language, this rate function is a good rate function in the sense that its level sets $\{\mathcal{I} \leq a\}$, $a \geq 0$, are compact in E (due to the compactness of the $\mathcal{H}$-balls in E).

2 The large deviation principle

Large deviations for Gaussian measures go back to M. Schilder [12] for the Wiener measure, and to M. Donsker and S. Varadhan [9] in general. The study of [9] actually adresses the large deviation principle for sums of independent Banach space valued random variables, the Gaussian case being a particular case.

In the context exposed in the first section, the following theorem presents the large deviation behavior of the law of εX as $\varepsilon \to 0$. For a subset A of E, let

$$\mathcal{I}(A) = \inf_{x \in A} \mathcal{I}(x).$$

Theorem 1 (The Gaussian large deviation principle). *For any closed set F in E,*

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in F) \leq -\mathcal{I}(F). \tag{4}$$

For any open set O in E,

$$\liminf_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in O) \ \geq \ -\mathcal{I}(O). \tag{5}$$

Applied to complements of balls, this theorem easily produces the limit (1), together with the observation that $\sigma = \sup_{|h|_{\mathcal{H}} \leq 1} \|h\|$.

The proof of the upper-bound (4) in Theorem 1 presented here relies on isoperimetric and concentration inequalities (cf. [3, 4]) which provide a very convenient tool to this task. The lower-bound (5) classically relies on the Cameron-Martin translation formula. The combined arguments actually produce a measurable version of the large deviation principle, without referring to any topology associated to the underlying abstract Wiener space (cf. [5, 10]).

Proof. A simple proof of the upper-bound (4) may therefore be provided by the Gaussian isoperimetric inequality (actually only the suitable concentration properties). Namely, let F be closed in E, and take r such that $0 < r < \mathcal{I}(F)$. By the very definition of $\mathcal{I}(F)$,

$$F \cap \sqrt{2r}\, \mathcal{K} \ = \ \emptyset,$$

where $\mathcal{K}$ is the (closed) unit ball in $\mathcal{H}$. Since F is closed and $\mathcal{K}$ is compact in E, there exists $\eta > 0$ such that it still holds true that

$$F \cap \big[\sqrt{2r}\, \mathcal{K} + B_E(0, \eta) \big] \ = \ \emptyset$$

where $B_E(0, \eta)$ is the ball with center the origin and with radius η for the norm $\| \cdot \|$ in E. Clearly

$$\lim_{\varepsilon \to 0} \mathbb{P}\big(\varepsilon X \in B_E(0, \eta)\big) = \lim_{\varepsilon \to 0} \mathbb{P}\big(X \in B_E(0, \tfrac{\eta}{\varepsilon})\big) \ = \ 1.$$

Recall now the Gaussian isoperimetric inequality for the law of X (cf. [3]), expressing that, whenever $\mathbb{P}(X \in A) \geq \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{1}{2}x^2} dx$ for some $a \in \mathbb{R}$,

$$\mathbb{P}(X \in A + s\, \mathcal{K}) \ \geq \ \Phi(a + s)$$

for every $s \geq 0$. For $\varepsilon > 0$ small enough, $\mathbb{P}(X \in B_E(0, \frac{\eta}{\varepsilon})) \geq \frac{1}{2} = \Phi(0)$. Hence,

$$\mathbb{P}(\varepsilon X \in F) \ \leq \ \mathbb{P}\big(\varepsilon X \notin \sqrt{2r}\, \mathcal{K} + B_E(0, \eta)\big) \ \leq \ 1 - \Phi\bigg(\frac{\sqrt{2r}}{\varepsilon}\bigg) \ \leq \ e^{-r/\varepsilon^2}.$$

Therefore

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in F) \ \leq \ -r,$$

which is the result since $r < \mathcal{I}(F)$ is arbitrary.

As mentioned above, the full strength of the Gaussian isoperimetric inequality is not really needed, and weaker concentration inequalities are enough to achieve the conclusion. For example, as emphasized in [4],

$$\mathbb{P}(X \in A + s\,\mathcal{K}) \,\geq\, 1 - e^{-\frac{1}{2}s^2 + \delta(\mu(A))s}$$

for every $s \geq 0$, where $\delta(\mu(A)) \to 0$ as $\mu(A) \to 1$, so that the proof may be developed similarly.

The proof of the lower-bound (5) is an application of the Cameron-Martin translation formula. Let $h \in O \cap \mathcal{H}$. Since O is open, there exists $\eta > 0$ such that $h + B_E(0, \eta) \subset O$, and thus

$$\mathbb{P}(\varepsilon X \in O) \,\geq\, \mathbb{P}\big(\varepsilon X \in h + B_E(0, \eta)\big).$$

In the notation of [2], the Cameron-Martin translation formula yields that

$$\begin{aligned} \mathbb{P}\big(\varepsilon X \in h + B_E(0, \eta)\big) \,&=\, \mu\big(\tfrac{h}{\varepsilon} + B_E(0, \tfrac{\eta}{\varepsilon})\big) \\ &=\, \exp\bigg(-\frac{|h|_{\mathcal{H}}^2}{2\varepsilon^2}\bigg) \int_{B_E(0, \frac{\eta}{\varepsilon})} \exp\bigg(-\frac{\widetilde{h}}{\varepsilon}\bigg) d\mu, \end{aligned}$$

where it is recalled that $\widetilde{h}$ is Gaussian under μ with variance $|h|_{\mathcal{H}}^2$ ($\widetilde{h} = \int_0^1 h'(t)dW(t)$ on the Wiener space). By Jensen's inequality,

$$\int_{B_E(0, \frac{\eta}{\varepsilon})} \exp\bigg(-\frac{\widetilde{h}}{\varepsilon}\bigg) d\mu \,\geq\, \mu\big(B_E(0, \tfrac{\eta}{\varepsilon})\big) \exp\bigg(-\int_{B_E(0, \frac{\eta}{\varepsilon})} \frac{\widetilde{h}}{\varepsilon} \cdot \frac{d\mu}{\mu(B_E(0, \frac{\eta}{\varepsilon}))}\bigg).$$

Now

$$\int_{B_E(0, \frac{\eta}{\varepsilon})} \widetilde{h}\, d\mu \,\leq\, \int_E |\widetilde{h}| d\mu \,\leq\, \bigg(\int_E \widetilde{h}^2 d\mu\bigg)^{1/2} \,=\, |h|_{\mathcal{H}}.$$

For every $\varepsilon > 0$ small enough, $\mu(B_E(0, \frac{\eta}{\varepsilon})) \geq \frac{1}{2}$ (for example). As a consequence of the various preceding lower-bounds,

$$\mathbb{P}(\varepsilon X \in O) \,\geq\, \frac{1}{2} \exp\bigg(-\frac{|h|_{\mathcal{H}}^2}{2\varepsilon^2} - \frac{2|h|_{\mathcal{H}}}{\varepsilon}\bigg)$$

from which it follows that

$$\liminf_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in O) \,\geq\, -\frac{1}{2}\, |h|_{\mathcal{H}}^2 \,=\, -\mathcal{I}(h).$$

This result for any $h \in O \cap \mathcal{H}$ yields the announced lower-bound (5), and completing therefore the proof of Theorem 1. □

References

[1] Integrability of norms of Gaussian random vectors and processes. *The Gaussian Blog.*

[2] Admissible shift, reproducing kernel Hilbert space, and abstract Wiener space. *The Gaussian Blog.*

[3] The Gaussian isoperimetric inequality. *The Gaussian Blog.*

[4] Gaussian concentration inequalities. *The Gaussian Blog.*

[5] G. Ben Arous, M. Ledoux. Schilder's large deviation principle without topology. *Pitman Res. Notes Math. Ser.* 284, 107–121 (1993).

[6] V. Bogachev. *Gaussian measures.* Math. Surveys Monogr. 62. American Mathematical Society (1998).

[7] A. Dembo, O. Zeitouni. *Large Deviations Techniques and Applications.* Springer (1998).

[8] J.-D. Deuschel, D. Stroock. *Large deviations.* Academic Press (1989).

[9] M. D. Donsker, S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time III. *Comm. Pure Appl. Math.* 29, 389–461 (1976).

[10] M. Ledoux. Isoperimetry and Gaussian Analysis. *École d'Été de Probabilités de St-Flour 1994. Lecture Notes in Math.* 1648, 165–294. Springer (1996).

[11] M. Lifshits. *Gaussian random functions.* Kluwer (1994).

[12] M. Schilder. Asymptotic formulas for Wiener integrals. *Trans. Amer. Math. Soc.* 125, 63–85 (1966).

[13] D. Stroock. *Probability theory: an analytic view.* Cambridge University Press (2011).

[14] S. R. S. Varadhan. *Large Deviations and Applications.* SIAM Publications, Philadelphia (1984).