Gaussian comparison inequalities

The law of a centered Gaussian random vector $X = (X_1, \ldots, X_n)$, on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with values in $\mathbb{R}^n$, is uniquely determined by its covariance matrix Σ^X with entries

$$\Sigma^X_{k\ell} = \mathbb{E}(X_k X_\ell), \quad k, \ell = 1, \ldots, n.$$

Given then two such Gaussian vectors $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ such that the respective covariance matrices Σ^X and Σ^Y may be compared in some way, it is expected that some statistics of the samples $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ may also be also be compared.

A famous result in this regard is the Slepian inequality [24] (see also [23, 25, 12, 16]) which expresses that, provided

$$\Sigma^X_{k\ell} = \mathbb{E}(X_k X_\ell) \leq \mathbb{E}(Y_k Y_\ell) = \Sigma^Y_{k\ell} \quad \text{for all } k, \ell = 1, \ldots, n,$$

with equality on the diagonal $k = \ell$, then, for every $u_1, \ldots, u_n \in \mathbb{R}$,

$$\mathbb{P}\bigg(\bigcup_{k=1}^n \{Y_k > u_k\} \bigg) \leq \mathbb{P}\bigg(\bigcup_{k=1}^n \{X_k > u_k\} \bigg). \tag{1}$$

In particular, by integration by parts,

$$\mathbb{E}\big(\max_{1 \leq \leq n} Y_k \big) \leq \mathbb{E}\big(\max_{1 \leq k \leq n} X_k \big).$$

This result is part of the comparison properties between Gaussian vectors. The post illustrates some of them, including the Anderson inequality, the Sudakov-Chevet-Fernique

Drafted by M. L. **v1** January 2024

inequality, and the Gordon min-max inequality. General references on the topic include [11, 8, 13, 28, 5, 20, 21, 14, 9, 22] etc.

Table of contents

1 Anderson's inequality

Anderson's inequality [7] is a useful comparison property which holds true under the stronger hypothesis that the matrix $\Sigma^Y - \Sigma^X$ is positive-definite.

Theorem 1 (Anderson's inequality)**.** *Let* $X = (X_1, \ldots, X_n)$ *and* $Y = (Y_1, \ldots, Y_n)$ *be two centered Gaussian random vectors in* $\mathbb{R}^n$ *such that*

$$\mathbb{E}\big(\langle c, X\rangle^2\big) \,\leq\, \mathbb{E}\big(\langle c, Y\rangle^2\big)$$

for any $c \in \mathbb{R}^n$*. Then, for any convex symmetric set* C *in* $\mathbb{R}^n$*,*

$$\mathbb{P}(Y \in C) \,\leq\, \mathbb{P}(X \in C). \tag{2}$$

Proof. Let Z be centered Gaussian with covariance $\Sigma^Y - \Sigma^X$, which is positive-definite by the hypothesis, and independent from X. By construction, Y has the same law as $X + Z$. By Fubini's theorem

$$\mathbb{P}(Y \in C) \,=\, \mathbb{P}(X + Z \in C) \,=\, \int_{\mathbb{R}^n} \mathbb{P}(X + z \in C) d\,\mathbb{P}_Z(z)$$

with $\mathbb{P}_Z$ the law of Z. The log-concavity property of Gaussian measures (cf. [1]) implies that, for every $z \in \mathbb{R}^n$,

$$\begin{aligned} \log \mathbb{P}(X \in C) \,&=\, \log \mathbb{P}\big(X \in \tfrac{1}{2}(C + z) + \tfrac{1}{2}(C - z)\big) \\ &\geq\, \frac{1}{2} \log \mathbb{P}(X \in C + z) + \frac{1}{2} \log \mathbb{P}(X \in C - z). \end{aligned}$$

Therefore, by symmetry of C (and of X), $\mathbb{P}(X \in C) \geq \mathbb{P}(X + z \in C)$, and the conclusion follows. □

It should be noticed that, without the symmetry assumption on C, it still holds true that

$$\mathbb{P}(X \notin C) \,\leq\, 2\,\mathbb{P}(Y \notin C).$$

Indeed, by convexity of C,

$$\begin{aligned}\mathbb{P}(X \notin C) &= \mathbb{P}\big(\tfrac{1}{2}\,(X+Z) + \tfrac{1}{2}\,(X-Z) \notin C\big) \\ &\leq \mathbb{P}\big(\{X+Z \notin C\} \cup \{X-Z \notin C\}\big) \\ &\leq \mathbb{P}(X+Z \notin C) + \mathbb{P}(X-Z \notin C),\end{aligned}$$

and the claim follows since Y has the same law as $X+Z$ as well as $X-Z$.

2 The Sudakov-Chevet-Fernique inequality

Slepian's inequality from the introduction is a most useful tool in applications, but however requires equality on the diagonal of the covariance matrices, which might be inconvenient.

Another version of the Slepian inequality compares rather the respective L^2-moments of the increments. This formulation, going back to V. Sudakov, S. Chevet and X. Fernique [26, 8, 13, 6], is in particular most efficient in the study of lower-bounds on suprema of Gaussian processes (cf. [2]).

Theorem 2 (The Sudakov-Chevet-Fernique inequality)**.** *Let $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ be two centered Gaussian random vectors in $\mathbb{R}^n$ such that*

$$\mathbb{E}\big([Y_k - Y_\ell]^2\big) \,\leq\, \mathbb{E}\big([X_k - X_\ell]^2\big) \quad \textit{for all} \quad k, \ell = 1, \ldots, n.$$

Then

$$\mathbb{E}\big(\max_{1\leq k\leq n} Y_k\big) \,\leq\, \mathbb{E}\big(\max_{1\leq k\leq n} X_k\big).$$

This theorem has been extended in various directions. It is shown for example in [13] that, under the same hypotheses, for any positive convex increasing function f on $\mathbb{R}_+$,

$$\mathbb{E}\Big(f\Big(\max_{1\leq k,\ell\leq n} [Y_k - Y_\ell]\Big)\Big) \,\leq\, \mathbb{E}\Big(f\Big(\max_{1\leq k,\ell\leq n} [X_k - X_\ell]\Big)\Big).$$

The proof of Theorem 2, as well as of the Slepian inequality (1), relies on a basic interpolation scheme, in the spirit of the Ornstein-Uhlenbeck semigroup [3]. At a technical level, it requires a suitable smoothing of the maximum function, which has been addressed in various ways in the literature. While the early proofs involved the use of derivative in the distributional sense, alternate arguments have been produced later. The proof displayed below uses ideas from statistical mechanics emphasized in [10]. The centering hypothesis is lifted by the condition $\mathbb{E}(X) = \mathbb{E}(Y)$ in [29].

Proof. It may be assumed that X and Y are independent. For any $t \in [0,1]$, set

$$Z_t \,=\, \sqrt{t}\, X + \sqrt{1-t}\, Y$$

(which thus linearly interpolates between X and Y). For any $\beta > 0$, let

$$F_\beta(x) \,=\, \frac{1}{\beta} \log\bigg(\sum_{k=1}^n e^{\beta x_k} \bigg), \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

The goal will be to show that the function

$$\phi(t) \,=\, \mathbb{E}\big(F_\beta(Z_t)\big), \quad t \in [0,1],$$

is increasing. The function ϕ is differentiable on $(0,1)$ with

$$\phi'(t) \,=\, \frac{1}{2} \sum_{k=1}^n \mathbb{E}\bigg(\bigg(\frac{X_k}{\sqrt{t}} - \frac{Y_k}{\sqrt{1-t}} \bigg) \partial_k F_\beta(Z_t) \bigg).$$

Now, by the integration by parts formula for Gaussian vectors (cf. [4]), for every $k = 1, \ldots, n$,

$$\mathbb{E}\big(X_k\, \partial_k F_\beta(Z_t)\big) \,=\, \sqrt{t} \sum_{\ell=1}^n \Sigma^X_{k\ell}\, \mathbb{E}\big(\partial_{k\ell} F_\beta(Z_t)\big)$$

and similarly for $\mathbb{E}(Y_k\, \partial_k F_\beta(Z_t))$. Combining the two,

$$\phi'(t) \,=\, \frac{1}{2} \sum_{k,\ell=1}^n (\Sigma^X_{k\ell} - \Sigma^Y_{k\ell})\, \mathbb{E}\big(\partial_{k\ell} F_\beta(Z_t)\big). \tag{3}$$

Next, at every $x \in \mathbb{R}^n$,

$$\partial_k F_\beta(x) \,=\, p_k(x) \,=\, \frac{e^{\beta x_k}}{\sum_{\ell=1}^n e^{\beta x_\ell}}\,, \quad k = 1, \ldots, n,$$

and

$$\partial_{k\ell} F_\beta(x) \,=\, \begin{cases} \beta(p_k(x) - p_k(x)^2) & \text{if} \;\; k = \ell, \\ -\beta\, p_k(x) p_\ell(x) & \text{if} \;\; k \neq \ell. \end{cases}$$

Hence

$$\sum_{k,\ell=1}^n (\Sigma^X_{k\ell} - \Sigma^Y_{k\ell})\, \partial_{k\ell} F_\beta(x) \,=\, \beta \sum_{k=1}^n (\Sigma^X_{kk} - \Sigma^Y_{kk})\, p_k(x) - \beta \sum_{k,\ell=1}^n (\Sigma^X_{k\ell} - \Sigma^Y_{k\ell})\, p_k(x) p_\ell(x)$$

which can be rewritten, using that $\sum_{k=1}^n p_k(x) = 1$, as

$$\sum_{k,\ell=1}^n (\Sigma^X_{k\ell} - \Sigma^Y_{k\ell})\, \partial_{k\ell} F_\beta(x) \,=\, \frac{\beta}{2} \sum_{k,\ell=1}^n \Big(\mathbb{E}\big([X_k - X_\ell]^2\big) - \mathbb{E}\big([Y_k - Y_\ell]^2\big) \Big) p_k(x) p_\ell(x).$$

Therefore, by the hypothesis and (3), $\phi'(t) \geq 0$ from which

$$\mathbb{E}\big(F_\beta(Y)\big) \;=\; \phi(0) \;\leq\; \phi(1) \;=\; \mathbb{E}\big(F_\beta(X)\big).$$

It remains to observe that $\lim_{\beta\to\infty} F_\beta(x) = \max_{1\leq k\leq n} x_k$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

□

3 The Gordon min-max inequalities

The Gordon min-max inequality [15] is an extension of Slepian's inequality to arrays of (jointly) Gaussian random variables. Further developments appear in [18, 17].

Theorem 3 (The Gordon min-max inequality I)**.** *Let $X = (X_{ij})_{1\leq i\leq n, 1\leq j\leq m}$ and $Y = (Y_{ij})_{1\leq i\leq n, 1\leq j\leq m}$ be two centered Gaussian random vectors in $\mathbb{R}^{nm}$ such that*

$$\begin{cases} \mathbb{E}(X_{ij}X_{ik}) \leq \mathbb{E}(Y_{ij}Y_{ik}) & \text{for all} \quad i, j, k, \\ \mathbb{E}(X_{ij}X_{\ell k}) \geq \mathbb{E}(Y_{ij}Y_{\ell k}) & \text{for all} \quad i \neq \ell \text{ and all } j, k \\ \mathbb{E}(X_{ij}^2) = \mathbb{E}(Y_{ij}^2) & \text{for all} \quad i, j. \end{cases}$$

Then, for every family $(u_{ij})_{1\leq i\leq n, 1\leq j\leq m}$ of real numbers,

$$\mathbb{P}\bigg(\bigcap_{i=1}^{n}\bigcup_{j=1}^{m}\{Y_{ij} > u_{ij}\}\bigg) \;\leq\; \mathbb{P}\bigg(\bigcap_{i=1}^{n}\bigcup_{j=1}^{m}\{X_{ij} > u_{ij}\}\bigg).$$

In particular,

$$\mathbb{E}\big(\min_{1\leq i\leq n}\max_{1\leq j\leq m} Y_{ij}\big) \;\leq\; \mathbb{E}\big(\min_{1\leq i\leq n}\max_{1\leq j\leq m} X_{ij}\big).$$

There is a similar version in terms of L^2-distances.

Theorem 4 (The Gordon min-max inequality II)**.** *Let $X = (X_{ij})_{1\leq i\leq n, 1\leq j\leq m}$ and $Y = (Y_{ij})_{1\leq i\leq n, 1\leq j\leq m}$ be two centered Gaussian random vectors in $\mathbb{R}^{nm}$ such that*

$$\begin{cases} \mathbb{E}\big(|Y_{ij} - Y_{ik}|^2\big) \leq \mathbb{E}\big(|X_{ij} - X_{ik}|^2\big) & \text{for all} \quad i, j, k, \\ \mathbb{E}\big(|Y_{ij} - Y_{\ell k}|^2\big) \geq \mathbb{E}\big(|X_{ij} - X_{\ell k}|^2\big) & \text{for all} \quad i \neq \ell \text{ and all } j, k. \end{cases}$$

Then

$$\mathbb{E}\big(\min_{1\leq i\leq n}\max_{1\leq j\leq m} Y_{ij}\big) \;\leq\; \mathbb{E}\big(\min_{1\leq i\leq n}\max_{1\leq j\leq m} X_{ij}\big).$$

References

[1] Some geometric inequalities for Gaussian measures. *The Gaussian blog*.

[2] Boundedness and continuity of Gaussian processes. *The Gaussian blog*.

[3] Mehler kernel, and the Ornstein-Uhlenbeck operator. *The Gaussian blog*.

[4] Some basics on Gaussian measures and variables. *The Gaussian blog*.

[5] R. Adler. *An introduction to continuity, extrema, and related topics for general Gaussian processes*. Lecture Notes–Monograph Series 12. Institute of Mathematical Statistics (1990).

[6] R. Alexander. Lipschitzian mappings and total mean curvature of polyhedral surfaces I. *Trans. Amer. Math. Soc.* 288, 661–678 (1985).

[7] T. Anderson: The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* 6, 170–176 (1955).

[8] A. Badrikrian, S. Chevet. *Mesures cylindriques, espaces de Wiener et fonctions aléatoires gaussiennes*. Lecture Notes in Math. 379. Springer (1974).

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

[10] S. Chatterjee. An error bound in the Sudakov-Fernique inequality (2008).

[11] R. Dudley. Sample functions of the Gaussian process. *Ann. Probab.* 1, 66–103 (1973).

[12] R. Dudley, D. Stroock. Slepian's inequality and commuting semigroups. *Séminaire de Probabilités XXI. Lecture Notes in Math.* 1247, 574–578. Springer (1987).

[13] X. Fernique. Régularité des trajectoires des fonctions aléatoires gaussiennes. *École d'Été de Probabilités de St-Flour 1974. Lecture Notes in Math.* 480, 1–96. Springer (1975).

[14] X. Fernique. *Fonctions aléatoires gaussiennes, vecteurs aléatoires gaussiens*. Université de Montréal, Centre de Recherches Mathématiques (1997).

[15] Y. Gordon. Some inequalities for Gaussian processes and applications. *Israel J. Math.* 50, 265–289 (1985).

[16] M. Gromov. Monotonicity of the volume of intersection of balls. *Geometric Aspects of Functional Analysis, Israel Seminar 1985/86. Lecture Notes in Math.* 1267, 1–4. Springer (1987)

[17] J. Hoffmann-Jørgensen. Slepian's inequality, modularity and integral orderings. *Progr. Probab.* 66, 19–53. Birkhäuser/Springer (2013).

[18] J. P. Kahane. Une inegalité du type de Slepian et Gordon sur les processus gaussiens. *Israel J. Math.* 55, 109–110 (1986).

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

[20] M. Ledoux, M. Talagrand. *Probability in Banach spaces (Isoperimetry and processes).* Ergebnisse der Mathematik und ihrer Grenzgebiete 23. Springer (1991).

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

[22] M. Lifshits. *Lectures on Gaussian processes.* SpringerBriefs Math. (2012).

[23] R. Plackett. A reduction formula for normal multivariate integrals. *Biometrika* 41, 351–360 (1954).

[24] D. Slepian. The one-sided barrier problem for Gaussian noise. *Bell. System Tech. J.* 41, 463–501 (1962).

[25] V. N. Sudakov. Gaussian measures, Cauchy measures and ε-entropy. *Soviet Math. Dokl.* 10, 310–313 (1969).

[26] V. N. Sudakov. Gaussian random processes and measures of solid angles in Hilbert spaces. *Soviet Math. Dokl.* 12, 412–415 (1971).

[27] V. N. Sudakov. A remark on the criterion of continuity of Gaussian sample functions. *Proceedings of the Second Japan-USSR Symposium on Probability Theory. Lecture Notes in Math.* 330, 444–454. Springer (1973).

[28] V. N. Sudakov. *Geometric problems in the theory of infinite-dimensional probability distributions (in Russian).* Trudy Mat. Inst. Steklov 141 (1976). English translation Proc. Steklov Inst. Math., American Mathematical Society (1979).

[29] R. Vitale. Some comparisons for Gaussian processes. *Proc. Amer. Math. Soc.* 128, 3043–3046 (2000).