A Festschrift for Herman Rubin
Institute of Mathematical Statistics
Lecture Notes – Monograph Series
Vol. 45 (2004) 337–341

A question of geometry and probability

Richard A. Vitale[1]

University of Connecticut

Abstract: We introduce the Aleksandrov–Fenchel inequality, apply it to a tail bound for Gaussian processes, and speculate on a further connection.

1. Introduction

Some time ago I brought a question involving geometry and probability to Herman and that led us in an interesting direction [12]. To celebrate this occasion, I am happy to bring another such question.

Recall the planar isoperimetric inequality, which says that for a convex body K of area $A(K)$ and perimeter $L(K)$

$$4\pi \cdot A(K) \leq L^2(K)\ . \tag{1}$$

Consider now a 2×2 matrix M of independent $N(0,1)$ variables and the image body MK. Inserting into (1) and taking expectations gives

$$4\pi \cdot E\left[A(MK)\right] \leq E\left[L^2(MK)\right] . \tag{2}$$

However, it is the case that the following stronger inequality holds:

$$4\pi \cdot E\left[A(MK)\right] \leq \left[EL(MK)\right]^2 . \tag{3}$$

It is possible to verify (3) as a simple exercise in Gaussian determinants, but one cannot say that this approach gives a satisfying explanation of what is really going on, for example, why (2) and (3) differ precisely by $\operatorname{Var} L(MK)$.

In fact, a deep theory is in the background, and the question of the title is to ask how it can be systematically exploited in this and other stochastic contexts. In the next sections, we briefly outline the theory and then turn to a specific question connected with Gaussian processes.

2. The Aleksandrov–Fenchel inequality

The bound (3) can be regarded as the first in an infinite sequence of inequalities, each of which is a stochastic formulation of the Aleksandrov–Fenchel (A–F) inequality in convex geometry. The A–F inequality is well–known to specialists as a powerful tool, having as implications the isoperimetric inequality (in all dimensions) and the Brunn–Minkowski inequality [8]. It has been successfully applied to problems in combinatorics as well as to the resolution of the van der Waerden permanent conjecture [6, 7, 14, 15, 16]. Interestingly, the original plan for the classic compilation [4] was to have a sequel based entirely on the A–F inequality. A closely

[1]Department of Statistics, University of Connecticut, Storrs, CT 06269, USA. e-mail: `r.vitale@uconn.edu`

Keywords and phrases: Aleksandrov–Fenchel inequality, Brunn–Minkowski inequality, convex body, Gaussian process, intrinsic volume, isoperimetric inequality, mean width, Steiner formula, tail bound.

AMS 2000 subject classifications: primary 52A40; secondary 52A05, 52A39, 60D05, 60G15.

related inequality on mixed discriminants [1, 2, 3] has found applications in stochastic settings. In view of this background, it is surprising that the A–F inequality itself has not found more applications in stochastic settings. One exception is questions in the theory of Gaussian processes, to which we turn in the next section.

A quick introduction to the A–F inequality goes as follows. It is part of *Brunn–Minkowski Theory* [13], which deals with the interaction between volume evaluation and vector addition of convex bodies (i.e., compact, convex subsets). For convex bodies $K_1, K_2, \ldots, K_n$ in $\mathbb{R}^d$ and positive coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n$,

$$\mathrm{vol}\,(\lambda_1 K_1 + \lambda_2 K_2 + \cdots + \lambda_n K_n) = \sum_{i_1, i_2, \cdots, i_d = 1}^{n} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_d} V(K_{i_1}, K_{i_2}, \ldots, K_{i_d}), \tag{4}$$

where, without loss of generality, the coefficients $V(\cdot)$ are taken to be symmetric in their arguments. The A–F inequality then asserts the following:

Theorem 1. *For convex bodies $K_1, K_2, \ldots, K_d$ in $\mathbb{R}^d$,*

$$V^2(K_1, K_2, K_3, \ldots, K_d) \geq V(K_1, K_1, K_3, \ldots, K_d)\; V(K_2, K_2, K_3, \ldots, K_d). \tag{5}$$

For the special case of a parallel body $K + \lambda B$ (B, the unit ball in $\mathbb{R}^d$), (4) simplifies to the Steiner formula [17]

$$\mathrm{vol}(K + \lambda B) = \sum_{i=0}^{d} \lambda^i \omega_i V_{d-i}(K), \tag{6}$$

where ω_i is the volume of the unit ball in $\mathbb{R}^i$ and the $V_i(K)$ are the *intrinsic volumes of* K ($V_0 \equiv 1$). Then (5) translates to the sequence $\{i!\, V_i(K)\}_{i=0}^{\infty}$ being log–concave:

$$(i!\, V_i(K))^2 \geq (i-1)! V_{i-1}(K) \;\cdot\; (i+1)! V_{i+1}(K) \qquad i = 1, 2, \ldots, d-1 \tag{7}$$

(elsewhere this property has been called *ultra–logconcavity of order* ∞ [10, 11]).

3. Intrinsic volumes and Gaussian processes

The theory of Gaussian processes has been heavily influenced by convex geometry [5, 9, 18, 22, 23]. Here we draw especially on [19, 20, 21, 22].

A popular approach to Gaussian processes is *canonical indexing*: suppose that $A \subseteq \mathbb{R}^d$ and that $Z = (Z_1, Z_2, \ldots, Z_d)$ are iid $N(0,1)$ variables. A canonically indexed Gaussian process $X_A = \{X_t, t \in A\}$ has the form $X_t = \sum_1^d t_i Z_i = \langle t, Z \rangle$ (this process evidently has "rank" no greater than d, but similar definitions can be made in Hilbert space for more general processes). If $A = K$, a convex body, then intrinsic volumes come into play. For $j = 1, 2, \ldots, d$ define the vector process $X_t^{j*} = \left(X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(j)} \right)$, where the components are independent copies of X_t. Further, define the (random) convex body $X_K^{j*} = \overline{\mathrm{conv}}\{X_t^{j*},\, t \in K\} \subseteq \mathbb{R}^j$. Then

$$V_j(K) = \frac{(2\pi)^{j/2}}{j!\,\omega_j} E\mathrm{vol}_j \left(X_K^{j*} \right). \qquad j = 1, 2, \ldots, d \tag{8}$$

The *Wills functional* is given by

$$W(K) = E\left[\exp\sup_{t\in K}\left(X_t - \frac{1}{2}EX_t^2\right)\right] \tag{9}$$

and has the generating function expansion

$$W(rK) = \sum_{j=0}^{d}\left(\frac{r}{\sqrt{2\pi}}\right)^j V_j(K). \tag{10}$$

An important consideration for Gaussian processes is "size," which is traditionally interpreted as $\sup_t X_t$. Tail probability bounds are of various types, and we illustrate the application of the preceding ideas to a sharpening of a bound in [22]. Fix K, and recall that the A–F inequality implies that $a_j = j!\,V_j(K)$ is a log-concave sequence:

$$\log a_j \le \log a_i + (\log a_{i+1} - \log a_i)(j-i),$$

which implies

$$V_j(K) \le \frac{i!\,V_i(K)}{j!}\left(\frac{(i+1)V_{i+1}(K)}{V_i(K)}\right)^{j-i}.$$

Substituting into (10) and summing $j = 0, \ldots, \infty$ yields

$$\begin{aligned} W(rK) &\le i!\,V_i(K)\left(\frac{V_i(K)}{(i+1)V_{i+1}(K)}\right)^i \exp\left[\frac{(i+1)V_{i+1}(K)r}{\sqrt{2\pi}V_i(K)}\right] \\ &\le \frac{i!\,V_i(K)}{(2\pi)^{i/2}m_i^i(K)}\exp\left[m_i(K)r\right], \end{aligned}$$

where

$$m_i(K) = \frac{i\,V_i(K)}{\sqrt{2\pi}V_{i-1}(K)}. \tag{11}$$

A straightforward application of Markov's inequality then provides the bound

$$P(\sup_t X_t \ge a) \le \inf_i\left\{\frac{i!\,V_i(K)}{(2\pi)^{i/2}m_i^i(K)}\exp\left[-\frac{(m_i(K)-a)^2}{2\sigma^2}\right]\right\},$$

where $a > 0$ and $\sigma^2 = \sup_{t\in K} EX_t^2 = \sup_{t\in K}\|t\|^2$.

This brings us to the issue mentioned in the introduction: the way in which the values $m_i(K)$ have arisen suggests that they may be natural parameters of the process for other questions as well. It is easy to verify that $m_1(K)$ is at once $E\sup_{t\in K} X_t$ and proportional to the mean width of K and, as such, has linear dimension. The succeeding m_i also have linear units, and evidently provide alternate size measures for both K and $\{X_t, t \in K\}$. Their asymptotic behavior reflects the regularity of the process (see [24] for details), but it seems clear that their specific values must also calibrate successive i-th order properties of some type for the process. What these are remains for investigation.

References

[1] Aleksandrov, A. D. (1938). Zur Theorie der gemischten Volumina von konvexen Köpern IV (Russian, German summary). *Mat. Sbornik* **3**, 227–251.

[2] Bapat, R. B. (1988). Discrete multivariate distributions and generalized log-concavity. *Sankhyā Ser. A* **50**, 98–110.

[3] Bapat, R. B. (1990). Permanents in probability and statistics. *Linear Algebra Appl.* **127**, 3–25.

[4] Beckenbach, E. F., and Bellman, R. (1971). *Inequalities.* Springer-Verlag, New York.

[5] Borell, C. (1975). The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30**, 207–216.

[6] Egorychev, G. P. (1981). The solution of van der Waerden's problem for permanents. *Adv. in Math.* **42**, 299–305.

[7] Falikman, D. I. (1981). Proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix. (Russian) *Mat. Zametki* **29**, 931–938.

[8] Gardner, R. J. (2002). The Brunn-Minkowski inequality. *Bull. Amer. Math. Soc.* (N.S.) **39**, no. 3, 355–405.

[9] Landau, H. J., and Shepp, L. A. (1970). On the supremum of a Gaussian process. *Sankhyā Ser. A* **32**, 369–378.

[10] Liggett, T. M. (1997). Ultra logconcave sequences and negative dependence. *J. Combin. Theory Ser. A* **79**, 315–325.

[11] Pemantle, R. (2000). Towards a theory of negative dependence. Probabilistic techniques in equilibrium and nonequilibrium statistical physics. *J. Math. Phys.* **41**, 1371–1390. MR1757964

[12] Rubin, H. and Vitale, R. A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* **6**, 165-170.

[13] Schneider, R. (1993). *Convex Bodies: the Brunn-Minkowski Theory.* Camb. Univ. Press, New York.

[14] Stanley, R. P. (1981). Two combinatorial applications of the Aleksandrov-Fenchel inequalities. *J. Combin. Theory Ser. A* **31**, 56–65.

[15] Stanley, R. P. (1986). Two poset polytopes. *Discrete Comput. Geom.* **1**, 9–23.

[16] Stanley, R. P. (1989). Log-concave and unimodal sequences in algebra, combinatorics, and geometry. Graph theory and its applications: East and West (Jinan, 1986), 500–535, *Ann. New York Acad. Sci.*, **576**, New York Acad. Sci.

[17] Steiner, J. (1840). Von dem Krümmunsschwerpunkte ebener Curven. *J. Reine Angew. Math.* **21**, 33–63; *Ges. Werke*, vol. 2, Reimer, Berlin, 1882, 99–159.

[18] Sudakov, V. N., and Tsirel'son, B. S. (1978). Extremal properties of half-spaces for spherically invariant measures. *J. Soviet Math.* **9**, 9–18; translated from *Zap. Nauch. Sem. L.O.M.I.* **41**, 14–24 (1974). MR365680

[19] Tsirel'son, B. S. (1982). A geometric approach to maximum likelihood estimation for infinite–dimensional Gaussian location I. *Theory Prob. Appl.* **27**, 411–418.

[20] Tsirel'son, B. S. (1985). A geometric approach to maximum likelihood estimation for infinite–dimensional Gaussian location II. *Theory Prob. Appl.* **30**, 820–828.

[21] Tsirel'son, B. S. (1986). A geometric approach to maximum likelihood estimation for infinite–dimensional location III. *Theory Prob. Appl.* **31**, 470–483.

[22] Vitale, R. A. (1996). The Wills functional and Gaussian processes. *Ann. Probab.* **24**, 2172–2178.

[23] Vitale, R. A. (1999). A log–concavity proof for a Gaussian exponential bound. In *Contemporary Math.: Advances in Stochastic Inequalities* (T.P. Hill, C. Houdré, eds.) **234**, AMS, 209–212. MR1694774

[24] Vitale, R. A. (2001). Intrinsic volumes and Gaussian processes. *Adv. Appl. Prob.* **33**, 354–364.