Isodiametry and geometric variance bounds

Tongseok Lim & Robert McCann

University of Toronto

www.math.toronto.edu/mccann

7 December 2019

Two equivalent questions:

Q1: Suppose all realizations of a random vector $X(\omega)$ lie in a subset of $\mathbf{R}^n$ having diameter one. GIve a sharp bound for the variance of X.

Q2: Among (Borel) probability measures μ on $\mathbf{R}^n$ whose support has diameter one, which ones have the maximum variance?

Two equivalent questions:

Q1: Suppose all realizations of a random vector $X(\omega)$ lie in a subset of $\mathbf{R}^n$ having diameter one. GIve a sharp bound for the variance of X.

Q2: Among (Borel) probability measures μ on $\mathbf{R}^n$ whose support has diameter one, which ones have the maximum variance?

Theorem (Lim-McCann '19+)

The maximum is uniquely attained (apart from translations and rotations) by the measure which assigns mass $\frac{1}{n+1}$ to each vertex of a regular unit diameter simplex (i.e. an equilateral triangle if $n = 2$ or regular tetrahedron if $n = 3$.)

Example

Applications to pattern formation in flocking and swarming models from mathematical biology to be discussed 9am Monday in Monarchs Salon.

Several classical results:

Popoviciu '35: case $n = 1$; in this case the diameter constraint is convex, however.

Bhatia-Davis '00: if $X(\omega)$ takes values in $[a, A] \subset \mathbf{R}$ and has mean $\bar{x}$, then

$$\mathrm{Var}(X) \leq (A - \bar{x})(\bar{x} - a)$$

Several classical results:

Popoviciu '35: case $n = 1$; in this case the diameter constraint is convex, however.

Bhatia-Davis '00: if $X(\omega)$ takes values in $[a, A] \subset \mathbf{R}$ and has mean $\bar{x}$, then

$$\mathrm{Var}(X) \leq (A - \bar{x})(\bar{x} - a)$$

Our answer generalizes both, and turns out to be equivalent to:

Jung (1901): if $K \subset \mathbf{R}^n$ has diameter at most 1 then K lies in a ball of radius $r_n := \sqrt{\frac{n}{2n+2}}$. Moreover, K is either contained in a smaller ball or contains a regular, unit-diameter, simplex.

which is nowadays proved using a combinatorial result of Helly.

Generalizing the Bhatia-Davis' bound to random vectors

RECALL: the mean and variance of a probability measure μ on $\mathbf{R}^n$ are

$$\begin{aligned} \bar{x}(\mu) &:= \int_{\mathbf{R}^n} x d\mu(x) \\ \mathrm{Var}(\mu) &:= \int_{\mathbf{R}^n} |x - \bar{x}(\mu)|^2 d\mu(x) \\ &= -|\bar{x}(\mu)|^2 + \int_{\mathbf{R}^n} |x|^2 d\mu(x) \end{aligned}$$

DEFN: The convex dual to $\phi : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is its Legendre transform

$$\phi^*(y) := \sup_{x \in \mathbf{R}^n} y \cdot x - \phi(x).$$

Then $\phi^{**} =$ greatest convex lower-semicontinuous function less than ϕ.

DEFN: Given $K \subset \mathbf{R}^n$ compact define

$$\phi_K(x) := \begin{cases} -|x|^2 & \text{if } x \in K \\ +\infty & \text{else} \end{cases}$$

Proposition

Each probability measure μ on $K \subset \mathbf{R}^n$ with mean $\bar{x}$ satisfies

$$\mathrm{Var}(\mu) \leq -|\bar{x}|^2 - \phi_K^{**}(\bar{x});$$

equality holds iff μ is supported on a certain sphere enclosing K.

ex 1) (Bhatia and Davis '00) If $K = [a, A] \subset \mathbf{R}$ then

$$-\phi_K^{**}((1-t)a + t\bar{A}) = (1-t)a^2 + tA^2$$

hence

$$-\phi_K^{**}(x) = (A+a)x - Aa.$$

Example (Applications to sample geometries)

(a) (Ball) If $K = B_R(0)$ then $\mathrm{Var}(\mu) \leq R^2 - |\bar{x}(\mu)|^2$, and equality holds iff μ is supported on ∂K.

(b) (Ellipse) If $K = \{(x_1, x_2) \in \mathbf{R}^2 \mid (\frac{x_1}{a})^2 + (\frac{x_2}{b})^2 \leq 1\}$ with $a > b > 0$ and $\mathrm{Var}(\mu) = -|\bar{x}(\mu)|^2 - \phi_K^{**}(\bar{x}(\mu))$ then *spt* μ consists of at most two points.

(c) (Rectangular parallelopiped) If $K = \prod_{i=1}^n [-a_i, a_i]$ is non-empty, then $\mathrm{Var}(\mu) \leq -|\bar{x}(\mu)|^2 + \sum_{i=1}^n a_i^2$, and equality forces μ to be supported on the vertices of K.

(d) (Diamond) If $a_1 > a_2 > 0$ and $K = \{(x_1, x_2) \in \mathbf{R}^2 \mid |\frac{x_1}{a_1}| + |\frac{x_2}{a_2}| \leq 1\}$, then $\mathrm{Var}(\mu) \leq a_1^2 - \frac{a_1^2 - a_2^2}{a_2}|\bar{x}_2(\mu)| - |\bar{x}(\mu)|^2$ and equality forces μ to be concentrated at the two vertices of K farthest from the origin, plus at most one of its other two vertices.

Proof of PROP: linear programming duality

Let $\mathcal{P}_0(K)$ be the set of mean 0 probability measures on $K \subset \mathbf{R}^n$.
To maximize the linear function $\mathrm{Var}(\cdot)$ on the convex set $\mathcal{P}_0(K)$,
let $(h, p) \in \mathbf{R} \times \mathbf{R}^n$ be Lagrange multipliers for the mass and mean.

$$\begin{aligned}
& \sup_{\mu \in \mathcal{P}_0(K)} \int_K |x|^2 d\mu(x) \\
= & \sup_{\mu \in \mathcal{M}_+(K) := [0,\infty) \times \mathcal{P}(K)} \inf_{h \in \mathbf{R}, p \in \mathbf{R}^n} h\big(1 - \mu(K)\big) + \int_K \left(|x|^2 - 2p \cdot x\right) d\mu(x) \\
\leq & \inf_{p \in \mathbf{R}^n, h \in \mathbf{R}} \sup_{\mu \in \mathcal{M}_+(K)} \left[h + \int_K \left(|x|^2 - 2p \cdot x - h\right) d\mu(x) \right]
\end{aligned}$$

Proof of PROP: linear programming duality

Let $\mathcal{P}_0(K)$ be the set of mean 0 probability measures on $K \subset \mathbf{R}^n$.
To maximize the linear function $\mathrm{Var}(\cdot)$ on the convex set $\mathcal{P}_0(K)$,
let $(h, p) \in \mathbf{R} \times \mathbf{R}^n$ be Lagrange multipliers for the mass and mean.

$$\begin{aligned}
& \sup_{\mu \in \mathcal{P}_0(K)} \int_K |x|^2 d\mu(x) \\
= & \sup_{\mu \in \mathcal{M}_+(K) := [0,\infty) \times \mathcal{P}(K)} \inf_{h \in \mathbf{R}, p \in \mathbf{R}^n} h(1 - \mu(K)) + \int_K (|x|^2 - 2p \cdot x) d\mu(x) \\
\leq & \inf_{p \in \mathbf{R}^n, h \in \mathbf{R}} \sup_{\mu \in \mathcal{M}_+(K)} \left[h + \int_K (|x|^2 - 2p \cdot x - h) d\mu(x) \right] \\
= & \inf_{p \in \mathbf{R}^n} \inf_{h \geq |x|^2 - 2p \cdot x \ \forall x \in K} h \\
= & \inf_{p \in \mathbf{R}^n} \sup_{x \in K} |x|^2 - 2p \cdot x \quad = \quad \inf_{p \in \mathbf{R}^n} \phi_K^*(-2p) \\
= & -\phi_K^{**}(0)
\end{aligned}$$

Bilinearity (existence of a saddle) allows interchange of inf with sup!

... sphericity of support

Thus

$$\sup_{\mu \in \mathcal{P}_0(K)} \int_K \left(|x|^2 - 2p \cdot x\right) d\mu(x) \quad = \quad \inf_{p \in \mathbf{R}^n} \sup_{x \in K} |x|^2 - 2p \cdot x.$$

Choosing the optimal μ, $p \in \mathbf{R}^n$ and $x^* \in K$, setting $R = |x^* - p|$ and completing the square yields

$$|x - p|^2 \leq |x^* - p|^2 \qquad \forall x \in K,$$

... sphericity of support

Thus

$$\sup_{\mu \in \mathcal{P}_0(K)} \int_K \left(|x|^2 - 2p \cdot x\right) d\mu(x) \quad = \quad \inf_{p \in \mathbf{R}^n} \sup_{x \in K} |x|^2 - 2p \cdot x.$$

Choosing the optimal μ, $p \in \mathbf{R}^n$ and $x^* \in K$, setting $R = |x^* - p|$ and completing the square yields

$$|x - p|^2 \leq |x^* - p|^2 \qquad \forall x \in K,$$

i.e.,

$$K \subset \overline{B_R(p)}.$$

Moreover, equality holds μ-a.e., so μ vanishes outside $K \cap \partial B_R(p)$ as asserted. QED

Maximizing variance under diameter constraint only

- unit diameter constraint is nonconvex
- have a linear program for each $K \subset \mathbf{R}^n$ and $\bar{x} \in \mathbf{R}^n$
- may take $\bar{x} = 0$ and $K \subset \mathbf{R}^n$ compact and convex without loss
- the identity

$$\phi_{K-w}^{**}(0) = |w|^2 + \phi_K^{**}(w)$$

shows the attained variance bound $-\phi_K^{**}(0) \geq -\phi_{K-p}^{**}(0)$ maximized among translations $p \in \mathbf{R}^n$ iff ϕ_K^{**} (and hence ϕ_K^*) are minimized at the origin

Maximizing variance under diameter constraint only

- unit diameter constraint is nonconvex
- have a linear program for each $K \subset \mathbf{R}^n$ and $\bar{x} \in \mathbf{R}^n$
- may take $\bar{x} = 0$ and $K \subset \mathbf{R}^n$ compact and convex without loss
- the identity

$$\phi^{**}_{K-w}(0) = |w|^2 + \phi^{**}_K(w)$$

shows the attained variance bound $-\phi^{**}_K(0) \geq -\phi^{**}_{K-p}(0)$ maximized among translations $p \in \mathbf{R}^n$ iff ϕ^{**}_K (and hence ϕ^*_K) are minimized at the origin

- in this case the sphere mentioned previously is centered: $p = 0$

Lemma (Tension between centered mean and diameter constraints)

(a) If $K \subset \partial B_r(0)$ is a subset of the radius $r > r_n := \sqrt{\frac{n}{2n+2}}$ centered sphere in $\mathbf{R}^n$ and $\mathrm{diam}(K) \le 1$, then $0 \notin ConvexHull(K)$.

Lemma (Tension between centered mean and diameter constraints)

(a) If $K \subset \partial B_r(0)$ *is a subset of the radius* $r > r_n := \sqrt{\frac{n}{2n+2}}$ *centered sphere in* $\mathbf{R}^n$ *and* $\mathrm{diam}(K) \leq 1$*, then* $0 \notin$ *ConvexHull*(K)*.*
(b) If K *is a subset of the centered sphere in* $\mathbf{R}^n$ *of radius* r_n, $\mathrm{diam}(K) \leq 1$ *and* $0 \in$ *ConvexHull*(K)*, then* K *is the set of vertices of a unit n-simplex.*

Proof: Elementary geometric induction on dimension...

Lemma (Tension between centered mean and diameter constraints)

(a) If $K \subset \partial B_r(0)$ *is a subset of the radius* $r > r_n := \sqrt{\frac{n}{2n+2}}$ *centered sphere in* $\mathbf{R}^n$ *and* $\mathrm{diam}(K) \le 1$*, then* $0 \notin ConvexHull(K)$*.*
(b) If K *is a subset of the centered sphere in* $\mathbf{R}^n$ *of radius* r_n*,* $\mathrm{diam}(K) \le 1$ *and* $0 \in ConvexHull(K)$*, then* K *is the set of vertices of a unit n-simplex.*

Proof: Elementary geometric induction on dimension...

QED

Thank you