Geometry, analysis, and information theory

Deane Yang, Polytechnic University

11 May 2002

Outline of talk

1. Overview

2. A little information theory

   (a) Noisy signal transmission
   (b) Covariance and Fisher information
   (c) The Cramer–Rao inequality

3. Geometry

   (a) Brunn-Minkowski theory and its variants
   (b) The Cramer—Rao inequality for convex bodies

4. Isoperimetric and Sobolev inequalities

   (a) Euclidean invariants and inequalities
   (b) Affine and linear invariants and inequalities

5. Open problems

Current research (with Erwin Lutwak and Gaoyong Zhang)

Let X be a finite dimensional real vector space and $dx \in \Lambda^n X^* \backslash \{0\}$. We study sharp inequalities satisfied by linear and affine invariants of the following.

- (Geometry) Convex bodies $K \subset X$.

- (Analysis) Functions $f : X \to \mathbf{R}$.

- (Information theory) Probability distributions $f(x)\,dx$ on X.

A big question for others or the future: To what extent can the invariants and inequalities be extended to nonlinear spaces?

Connections between geometry, analysis, and information theory

- (Elliot Lieb) The Shannon entropy power inequality (information theory) and the Brunn–Minkowski inequality (geometry) both follow from the sharp Young inequality (analysis) proved by Beckner and Brascamp–Lieb.

- (Bill Beckner) The logarithmic Sobolev inequality of Gross is equivalent to an inequality due to Stam and Weissler involving the entropy and Fisher information of a probability distribution. Stam used this inequality to prove the Shannon entropy power inequality.

- (Szarek–Voiculescu) The Shannon entropy power inequality (information theory) follows from a restricted Brunn–Minkowski inequality (geometry).

There is an excellent survey article titled *The Brunn–Minkowski inequality* by Richard Gardner to appear in the Bulletin of the AMS.

Noisy transmission of a signal

- **Transmitted signal:** $x_0 \in X$, where X is an n–dimensional real vector space.
- **Received signal:** a random vector $x \in X$ with respect to a probability distribution

$$p(x - x_0)\, dx,$$

  where the distribution $p(x)\, dx$ has mean 0.
- **Single component of received signal:**

$$\ell(x) = \langle \xi, x \rangle$$

  where $\xi \in X^*$.
- **Mean square error:**

$$\int_X \langle \xi, x - x_0 \rangle^2 p(x - x_0)\, dx = C(\xi, \xi).$$

- **Covariance matrix:**

$$C = \int_X (x \otimes x) p(x)\, dx.$$

Mean estimation

- **Repeated transmission of the same signal:** Random vectors $x_1, \ldots, x_N \in X$.
- **Mean estimate:** The random vector

$$\overline{x} = \frac{x_1 + \cdots + x_N}{N}.$$

with distribution P_N.
- **The central limit theorem:** As $N \to \infty$,

$$P_N \to G\left(x_0, \frac{C}{\sqrt{N}}\right),$$

where G is the standard normal distribution.

Maximum likelihood estimation

- **Repeated transmission of the same signal:** Random vectors $x_1, \ldots, x_N \in X$.
- **Maximum likelihood estimate:** The random vector x_m that maximizes the *log likelihood function*,

$$\phi(x) = \sum_{i=1}^{N} \log p(x_i - x).$$

  with distribution Q_N.
- **Fisher information matrix:**

$$F = \int_{\mathbf{R}^n} (d(\log p(x)) \otimes d(\log p(x))) p \, dx.$$

- **According to Fisher and Doob:** As $N \to \infty$,

$$Q_N \to G\left(x_0, \frac{F^{-1}}{\sqrt{N}}\right)$$

The Cramer–Rao inequality

Theorem.

$$C \geq F^{-1}$$

with equality if and only if the distribution p *is Gaussian.*

In other words, the mean square error of the mean estimate is greater than or equal to the mean square error of the maximum likelihood estimate, with equality holding if and only if the distribution is Gaussian.

A geometric description of Cramer–Rao

- **Covariance ellipsoid**

$$E_C = \{x \in X \ : \ C^{-1}(x, x) \leq 1\}.$$

- **Fisher information ellipsoid**

$$E_F = \{x \in X \ : \ F(x, x) \leq 1\}.$$

- **Cramer–Rao inequality:**

$$E_F \subset E_C,$$

  with equality if and only if p is Gaussian.

Convex bodies

- A **convex body** is a compact convex set $K \subset X$ that contains the origin in its interior.

- A convex body $K \subset X$ is determined by its **support function** $h_K : X^* \to \mathbf{R}$, where
$$h_K(\xi) = \sup\{\langle \xi, x\rangle \ : \ x \in K\}.$$

- It is also determined by its **dual support function** $h^*_K : X \to \mathbf{R}$, where
$$h^*_K(x) = \inf\{\lambda \ : \ x/\lambda \in K\}.$$

Classical Brunn–Minkowski theory

- **Volume.** Fix a volume form dx on the vector space X. Let

$$V(K) = \int_K dx.$$

- **Minkowski addition:**

$$K + L = \{x + y \ : \ x \in K, \ y \in L\}.$$

  An equivalent form is

$$h_{K+L} = h_K + h_L.$$

- **Brunn–Minkowski inequality**

$$V(K + L)^{\frac{1}{n}} \geq V(K)^{\frac{1}{n}} + V(L)^{\frac{1}{n}},$$

  with equality holding if and only if K and L are equal up to translation and dilation.

- **Minkowski inequality** The *mixed volume* of K and L is defined to be

$$V(K,L) = \lim_{t\to 0+} \frac{V(K+tL) - V(K)}{t}.$$

Differentiating the Brunn–Minkowski inequality yields the Minkowski inequality,

$$V(K,L) \geq V(K)^{1-\frac{1}{n}} V(L)^{\frac{1}{n}},$$

with equality holding if and only if K and L are equal up to translation and dilation.

Remark. *Setting L equal to the unit Euclidean ball yields the classical Euclidean isoperimetric inequality.*

L^2 Brunn–Minkowski theory

- L^2 **Minkowski–Firey addition:**

$$(h_{K+_2L})^2 = (h_K)^2 + (h_L)^2$$

- L^2 **mixed volume:**

$$V_2(K,L) = \lim_{t\to 0+} \frac{V(K +_2 \sqrt{t}L) - V(K)}{t}.$$

- L^2 **dual Minkowski addition:**

$$K +_{-2} L = (K^* +_2 L^*)^*.$$

- L^2 **dual mixed volume:**

$$V_{-2}(K,L) = \lim_{t\to 0+} \frac{V(K +_{-2} t^{-1/2}L) - V(K)}{t}.$$

Covariance and Legendre ellipsoids

- **Covariance ellipsoid of a probability distribution $p(x)\,dx$:**

$$h_C(\xi)^2 = \int_X \langle \xi, x - \overline{x} \rangle^2 p(x - \overline{x})\,dx.$$

- **The Legendre ellipsoid $\Gamma_2 K$ of a convex body K:**

$$h_{\Gamma_2 K}(\xi)^2 = \frac{n+2}{V(K)} \int_K \langle \xi, x \rangle^2\,dx.$$

The ellipsoid $\Gamma_2 K$ is the unique ellipsoid such that

$$V(\Gamma_2 K) = \sup\{V(E) \,:\, V_{-2}(K, E) = V(K)\}.$$

Fisher information ellipsoids

- **Fisher information ellipsoid of a probability distribution:**

$$h^*_F(v)^2 = \int_X \langle d(\log p(x)), v\rangle^2 p(x)\, dx,$$

- **A new ellipsoid (Lutwak–Yang–Zhang):**

$$h^*_{\Gamma_{-2}K}(v)^2 = \frac{n}{V(K)} \int_K \langle dh_K(x), v\rangle^2 dx.$$

The ellipsoid $\Gamma_{-2}K$ is the unique ellipsoid such that

$$V(\Gamma_{-2}K) = \sup\{E \,:\, V_2(K,E) = V(K)\}.$$

Inequalities for the covariance ellipsoids

Theorem. *Any probability distribution $f(x)\,dx$ on X satisfies*

$$h[f] \leq \log[(2\pi e)^{\frac{n}{2}}\sqrt{\det C}],$$

where C is the covariance matrix. Equality holds if and only if f is a Gaussian distribution.

Theorem. [classical]

$$V(K) \leq V(\Gamma_2 K),$$

with equality if and only if K is an ellipsoid centered at the origin.

Inequalities for the Fisher information ellipsoids

Theorem. *Any probability distribution $f(x)\,dx$ on X satisfies*

$$h[f] \geq \log[(2\pi e)^{\frac{n}{2}} \det F^{-1}],$$

where F is the Fisher information matrix.

Theorem. [Lutwak–Yang–Zhang]

$$V(K) \geq V(\Gamma_{-2}K),$$

with equality if and only if K is an ellipsoid centered at the origin.

Remark. *The information theoretic inequality is a logarithmic Sobolev inequality first proved by Stam ($n = 1$) and Weissler ($n \geq 1$). Beckner–Pearson showed that it is equivalent to the logarithmic Sobolev inequality proved by Gross.*

Consequences

Theorem. [Milman–Pajor] *If $K \subset \mathbf{R}^n$ is an origin–symmetric convex body, then*

$$\frac{\sqrt{n+2}}{2\sqrt{3}} \frac{V(K)}{h_{\Gamma_2 K}(\xi)} \leq V(K \cap \xi^\perp) \leq \frac{n}{\sqrt{2(n+2)}} \frac{V(K)}{h_{\Gamma_2 K}(\xi)},$$

for all $\xi \in X^*$.

Theorem. [Lutwak–Yang–Zhang] *If $K \subset \mathbf{R}^n$ is an origin–symmetric convex body, then*

$$\frac{1}{2} V(K) h^*_{\Gamma_{-2} K}(v) \leq V(\pi_v(K)) \leq \frac{\sqrt{n}}{2} V(K) h^*_{\Gamma_{-2} K}(v),$$

for all $v \in X$.

Equality conditions: Equality holds on the lefthand side only if K is a right cylinder and u is orthogonal to its base. Equality holds on the righthand side only if K is a double right cone and u points along its axis.

A rough dictionary between information theory and L^2 Brunn–Minkowski theory

Information theory	L^2 BMt
entropy power	$V(K)$
covariance	$\Gamma_2 K$
Fisher information	$\Gamma_{-2} K$

The Cramer–Rao inequality

Theorem. [Cramer, Rao] *The covariance and Fisher information of a probability distribution $p(x)\,dx$ on X satisfies*

$$F^{-1} \leq C,$$

with equality if and only if the distribution is Gaussian.

Theorem. [Lutwak–Yang–Zhang] *Any convex body $K \subset X$ satisfies*

$$\Gamma_{-2}K \subset \Gamma_2 K,$$

with equality holding if and only if K is an ellipsoid centered at the origin.

The Cramer–Rao inequality is a real form of the uncertainty principle.

The Euclidean isoperimetric inequality

Fix an inner product on the vector space X, and let $dx \in \Lambda^n X^*$ be the naturally induced volume measure.

Given any compact set $K \subset X$, let ∂K denote its boundary and $V_{n-1}(\partial K)$ the $(n-1)$–dimensional Hausdorff measure of ∂K.

Theorem. *Given any Euclidean ball B and compact set $K \subset X$,*

$$\frac{V_{n-1}(\partial K)}{V(K)^{\frac{n-1}{n}}} \geq \frac{V_{n-1}(\partial B)}{V(B)^{\frac{n-1}{n}}},$$

with equality holding if and only if K is a Euclidean ball.

Easiest proof uses the Brunn–Minkowski inequality.

The sharp L^p Euclidean Sobolev inequality

Given $1 \leq p < n$, let

$$\phi(t) = \begin{cases} (1 + t^{\frac{p}{p-1}})^{\frac{n}{p}-1} & \text{, if } p > 1 \\ H(1-t) & \text{, if p = 1,} \end{cases}$$

where H is the Heaviside function.

Theorem. [Aubin, Talenti] *Given $1 \leq p < n$, there exists a constant $c(n,p)$ such that any function $f : X \to \mathbf{R}$ satisfies*

$$\|df\|_p \geq c(n,p)\|f\|_{s(n,p)},$$

where $s(n,p) = np/(n-p)$, and equality holds if and only if

$$f(x) = a\phi(b|x - x_0|),$$

for some $a \in \mathbf{R}$, $b \in (0,\infty)$, and $x_0 \in X$.

Proof of the sharp L^p Euclidean Sobolev inequality

- Use the coarea formula and the Euclidean isoperimetric inequality to prove the following fundamental rearrangement inequality,

$$\|df\|_p \geq \|d\hat{f}\|_p,$$

  with equality holding if and only if $f = \pm\hat{f}(x - x_0)$ for some $x_0 \in X$, where $\hat{f}$ is the rearrangement of f.

- It therefore suffices to prove the inequality for rotationally symmetric functions. This is a surprisingly difficult 1–dimensional problem in the calculus of variations.

- The Euclidean isoperimetric inequality and the sharp L^1 Euclidean Sobolev inequality are equivalent.

The polar projection body

Given a body $K \subset X$ and a nonzero vector $v \in X$, the $(n-1)$–dimensional volume of the projection of K in the direction v is given by the function

$$\pi_1(v) = \frac{c(n)}{V(K)} \int_{\partial K} |v \rfloor dx|.$$

Since π_1 is a convex function of v, it defines a convex body

$$\Pi_1^* K = \{x \ : \ \pi_1(x) \leq 1\}.$$

We call $\Pi_1^* K$ the *polar projection body.* The constant $c(n)$ is chosen so that $\Pi_1^* E = E$, for any ellipsoid E centered at the origin.

The volume of $\Pi_1^* K$ can be viewed as the average $(n-1)$–volume of all projections of K.

Reformulation

- Given a convex body $K \subset X$, there is a measurable map $\ell : \partial K \to \partial K^*$, such that
$$\langle \ell(x), x \rangle = 1,$$
for almost every $x \in \partial K$.

- It has a unique extension to a measurable map $\ell : X \to X^*$ homogeneous of degree 1.

- The map ℓ is called the *Legendre transform.*

- If K is convex, the coarea formula shows that

$$\pi_1(v) = \frac{c(n)}{V(K)} \int_K |\langle \ell(x), x \rangle| \, dx.$$

The L^p polar projection body

Lutwak observed that the definition of the polar projection body can be generalized as follows.

Definition. *Given $1 \leq p < \infty$, a convex body $K \subset X$, let*

$$\pi_p(v) = \frac{c(n,p)}{V(K)} \int_K |\langle \ell(x), x \rangle|^p \, dx.$$

The L^p polar projection body is the convex body given by

$$\Pi_p^* K = \{x \ : \ \pi_p(x) \leq 1\},$$

where the constant $c(n,p)$ is chosen so that $\Pi_p^* E = E$, for any ellipsoid centered at the origin.

Remark. *The L^p polar projection body of a convex body K is always an L^p zonoid and can be viewed as a natural smoothing of K, since*

$$\lim_{p \to \infty} \Pi_p^* K = K.$$

The sharp L^p Petty projection inequality

Theorem. [Petty ($p = 1$), Lutwak–Yang–Zhang ($p > 1$)] *Given* $1 \leq p < \infty$, *any ellipsoid* $E \subset X$ *centered at the origin and convex body* $K \subset X$,

$$V(\Pi_p^* K) \leq V(K)$$

with equality holding if and only if K *is an ellipsoid centered at the origin.*

Remark. *The* L^1 *ineqality is affine. The Euclidean isoperimetric inequality follows from the* L^1 *Petty projection inequality and the Hölder inequality.*

The sharp L^p affine Sobolev inequality

Theorem. [Zhang ($p = 1$), Lutwak–Yang–Zhang ($1 < p < n$)] *Given $1 \leq p < n$ and a function $f : X \to \mathbf{R}$,*

$$V(\Pi_p^* f)^{-\frac{1}{n}} \geq c(n,p)\|f\|_{s(n,p)}.$$

Equality holds if and only if

$$f(x) = a\phi(\sqrt{\langle x - x_0, A(x - x_0)\rangle}),$$

for some $a \in \mathbf{R}$ and positive definite $A \in S^2 X^*$.

Remark. *The proof of the L^p inequality requires both the L^1 and the L^p Petty projection inequalities. The sharp L^p Euclidean Sobolev inequality follows from the sharp L^p affine Sobolev inequality and the Hölder inequality.*

Open problem: The reverse Blaschke–Santalo inequality

We will say that a body $K \subset X$ is *centered at the origin*, if its center of mass is at the origin. The *polar body* K^* of K is given by

$$K^* = \{\xi \ : \ \langle \xi, x \rangle \leq 1, \ x \in K\}.$$

Theorem. [Blaschke–Santalo inequality] *Given any ellipsoid $E \subset X$ centered at the origin and convex body $K \subset X$,*

$$V(K)V(K^*) \leq V(E)V(E^*),$$

with equality holding if and only if K is an ellipsoid centered at the origin.

Conjecture. *Given any simplex $\Delta \subset X$ centered at the origin and convex body $K \subset X$ centered at the origin,*

$$V(K)V(K^*) \geq V(\Delta)V(\Delta^*),$$

with equality holding if and only if K is a simplex centered at the origin.

The L^p Petty conjecture: Geometric formulation

Conjecture. *Given $1 \leq p < n$, an ellipsoid $E \subset X$, and a convex body $K \subset X$, the L^p projection body $\Pi_p K$ of K satisfies*

$$V(\Pi_p K)V(K) \geq V(E)V(E^*)$$

If $p > 1$, equality holds only if K is an ellipsoid centered at the origin. If $p = 1$, equality holds only if K is an ellipsoid.

The L^p Petty conjecture: Analytic formulation

Conjecture. *Given* $1 \le p < n$, *Given functions* $f : X \to \mathbf{R}$ *and* $g : X^* \to \mathbf{R}$,

$$\left[\int_{X \times X^*} |\langle df(x), dg(\xi)\rangle|^p \, dx \, d\xi\right]^{\frac{1}{p}} \ge c(n,p)\|f\|_s\|g\|_s. \qquad (1)$$

Equality holds in (1) *if and only if*

$$f(x) = a\phi(\sqrt{\langle x - x_0, A(x - x_0)\rangle})$$
$$g(\xi) = \alpha\phi(\sqrt{\langle \xi - \xi_0, A^{-1}(\xi - \xi_0)\rangle}),$$

for some $a, \alpha \in \mathbf{R}$, $x_0 \in X$, $\xi_0 \in X^*$, *and positive definite* $A \in S^2X^*$.

Remark. *The conjecture is known for the case* $p = 2$. *The analytic form follows from the sharp* L^2 *Euclidean Sobolev inequality. The geometric form follows from the analytic form and also follows from an inequality proved by Lutwak–Yang–Zhang.*

The slicing problem

The *Legendre ellipsoid* $\Gamma_2 K \subset X$ of a convex body $K \subset X$ has support function $h_{\Gamma_2 K}$ given by

$$h_{\Gamma_2 K}(\xi)^2 = \frac{n+2}{V(K)} \int_K \langle \xi, x \rangle^2 \, dx$$

A classical result is that

$$V(\Gamma_2 K) \geq V(K),$$

with equality if and only if $K = \Gamma_2 K$.

The slicing problem

Conjecture. *There exists a positive constant* c *independent of the dimension* n *so that the following equivalent statements hold for any origin–symmetric convex body* $K \subset X$.

1. *Fix an inner product on* X. *Then*

$$V(K)^{\frac{n-1}{n}} \leq c[\sup\{V(K \cap \xi^{\perp}) \ : \ \xi \in X^*\backslash\{0\}\}].$$

2. *The Legendre ellipsoid* $\Gamma_2 K$ *of* K *satisfies*

$$V(\Gamma_2 K) \leq c^n V(K).$$

Open–ended Questions

- To what extent can the affine and linear invariants and inequalities be extended to nonlinear spaces?
- What is L^p information theory?
- The analytic inequalities seem more fundamental than the geometric inequalities. Shouldn't there be direct proofs of the sharp Sobolev inequalities that do not rely on geometric inequalities?
- Elucidate the affine and linear theory of convex bodies using symplectic and projective geometry.