Monotonicity of area in (non-symmetric) normed spaces

Tony Thompson

CIRM, Liminy

9–13 January 2012

OUTLINE

1. Set the scene.
2. Dual areas and volumes.
3. Normalizations.
4. Monotonicity and its consequences.
5. Non-symmetric norms.

1. Setting the scene

Finite dimensional real linear space $\mathbb{R}^n$.
Mostly concerned with $\mathbb{R}^2$ and $\mathbb{R}^3$.
This means:

- things are easy to visualize;
- it is easy to draw pictures;
- the words length, area and volume have (more or less) their usual connotations (precise definitions come later).

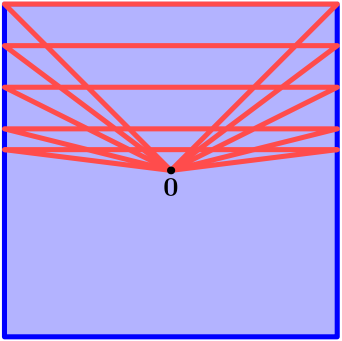
1. Setting the scene

This space is given a norm: $\|\cdot\|$
which makes it into both a metric space and a topological space
which have good properties.

- $d(t\mathbf{x}, t\mathbf{y}) = td(\mathbf{x}, \mathbf{y}) \quad (t \geq 0)$
  The metric scales properly.
- $d(\mathbf{x}+\mathbf{z}, \mathbf{y}+\mathbf{z}) = d(\mathbf{x}, \mathbf{y})$
  Translation invariant all translations are isometries.
- The topology is locally compact.

Area

The metric does not determine area in the plane. (Nor volume in $\mathbb{R}^3$.) For example:



All these triangles have sides $1, 1, 2$ in the ℓ_∞ metric but different areas. There is no formula of a Heron type.

Area

BUT

- because $\mathbb{R}^n$ is a locally compact group and
- because any sensible definition of area should be translation invariant
- any sensible area is a Haar measure — unique up to a scalar multiple.

The problem

- What is the appropriate normalization of Haar measure?
- “Appropriate” means “reflecting the geometry in some suitable sense”.
- For now, normalize by giving a suitable value for the area of the unit disc (volume of the unit ball).

2. Dual Measures

- Although Haar measure is not uniquely determined there is a unique measure on $\mathbb{R}^n \times (\mathbb{R}^n)^*$
- To see this, let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis for $\mathbb{R}^n$.
- The dual basis is given by $\{\mathbf{x}_1^*, \mathbf{x}_2^*, \ldots, \mathbf{x}_n^*\}$ where $\mathbf{x}_i^*(\mathbf{x}_j) = \delta_{ij}$.
- Let $P(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ be the parallelotope spanned by the basis vectors.

2. Dual measures

- If λ is a Haar measure in $\mathbb{R}^n$, normalize Haar measure λ^* in $(\mathbb{R}^n)^*$ so that
- $\lambda(P(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n))\lambda^*(P(\mathbf{x}_1^*, \mathbf{x}_2^*, \ldots, \mathbf{x}_n^*)) = 1.$
- λ^* is the measure dual to λ.

- Any other basis is of the form $\{T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n\}$ where T is an invertible linear map.
- Its dual is
$$\{T^{*-1}\mathbf{x}_1^*, T^{*-1}\mathbf{x}_2^*, \ldots, T^{*-1}\mathbf{x}_n^*\}$$
- because
$$(T^{*-1}\mathbf{x}_i^*)(\mathbf{x}_j) = \mathbf{x}_i^*(T^{-1}T\mathbf{x}_j) = \delta_{ij}.$$
- Then we also have:

$$\begin{aligned} \lambda(P(T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n)) &\times \lambda^*(P(T^{*-1}\mathbf{x}_1^*, T^{*-1}\mathbf{x}_2^*, \ldots, T^{*-1}\mathbf{x}_n^*)) \\ &= \det T(\det T)^{-1}.1 \\ &= 1. \end{aligned}$$

- Thus $\lambda \times \lambda^*$ on $\mathbb{R}^n \times (\mathbb{R}^n)^*$ is independent of basis;
- in particular

$$(t\lambda)^* = t^{-1}\lambda^*.$$

2. Dual measures

- If K is a closed, convex set containing $\mathbf{0}$ then:

$$K^* = \{\mathbf{p} : \mathbf{p}(\mathbf{x}) \leq 1 \;\; \forall \mathbf{x} \in K\}.$$

- If T is an invertible linear map $(TK)^* = T^{-1*}(K^*)$ because $(T^{-1*}\mathbf{p})(T\mathbf{x}) = \mathbf{p}(T^{-1}T\mathbf{x})$.
- If $\lambda \times \lambda^*$ is the Haar measure on $\mathbb{R}^n \times \mathbb{R}^{n*}$ then

$$\mathrm{vp}(K) := \lambda(K)\lambda^*(K^*) \tag{1}$$

is a linear invariant of K *i.e.*

$$\mathrm{vp}(T(K)) = \mathrm{vp}(K).$$

- Note that Equation (1) can be used to define λ^*.

3. Possible Normalizations.

- We normalize Haar measure by assigning a number $\mu(B)$ to be the area (in $\mathbb{R}^2$) of the unit ball B.
- If T is an invertible linear transformation, it is an isometry between the spaces $(\mathbb{R}^2, B)$ and $(\mathbb{R}^2, T(B))$. Hence, we should require $\mu(B) = \mu(T(B))$.
- In other words, the number we assign should be a linear invariant of the convex set B.

3. Possible Normalizations

- The simplest way to do this is to set $\mu_b(B) = \pi$; (the appropriate value in higher dimensions).
- However, once we make a normalization in $\mathbb{R}^2$ we automatically get a dual normalization in $(\mathbb{R}^2)^*$.
- This is obtained by setting $\mu(B)\mu^*(B^*) = \mathrm{vp}(B)$ (see equation (1)).
- In this case $\mu_b^*(B^*) = \mathrm{vp}(B)/\pi$.
- The roles can be reversed by setting $\mu_{ht}(B) = \mathrm{vp}(B)/\pi$. In other words, $\mu_b^* = \mu_{ht}$ and *vice versa*.

3. Possible Normalizations

- Once we have two (or more) possible normalizations it is possible to take convex combinations:

$$\mu(B) := t\mu_b(B) + (1-t)\mu_{ht}(B).$$

- Or other means, for example:

$$\mu_{sd}(B) = (\mu_b(B)\mu_{ht}(B))^{1/2} = \sqrt{\mathrm{vp}(B)}.$$

The subscript denotes that this definition is self-dual (and is the only one).

3. Possible Normalizations

- In $\mathbb{R}^2$ another linear invariant that one might choose is:

$$\mu_p(B) = \ell(\partial B)/2$$

  where ℓ denotes arc length (in the metric from B). (The subscript p denotes perimeter.)
- This number is, in general, hard to calculate.
- Having defined area one could, in theory, define volume in $\mathbb{R}^3$ by setting the volume of the ball to be $1/3$ its surface area. But this would present horrible computational problems!

3. Possible Normalizations

- Another method is to assign a value for the area of a convex set associated with B in a linearly invariant way.
  For example:
- If P is a smallest parallelogram (in higher dimensions parallelotope) circumscribing B set $\mu_{m*}(P) = 4$; (in higher dimensions 2^n).
  (It doesn't matter if P is not unique – it usually is not.)
- The choice of 4 (2^n) is to fit with the Euclidean case.
  The subscript $m*$ is because this is the normalization called mass* by Gromov.

The dual normalization (Gromov's mass) assigns the value 2 (in higher dimensions $2^n/n!$) to the largest parallelogram (cross-polytope) inscribed to B.
Instead of using parallelograms one can also use the (unique) largest inscribed and smallest circumscribed ellipses (ellipsoids) to B.
In either case we assign the area π (resp. ϵ_n) to this ellipse (resp. ellipsoid). Thus:

$$\mu_I(E) = \pi$$

where E is the minimal circumscribed ellipse to B.
The dual measure using the maximal inscribed one is denoted by μ_{I*}.

3. Possible Normalizations

- This gives us 8 (9 if we add the dual of μ_p) possibilities.
- The spaces between these definitions can be filled with convex combinations or other averages.
- How to sort out this morass?

3. Possible Normalizations

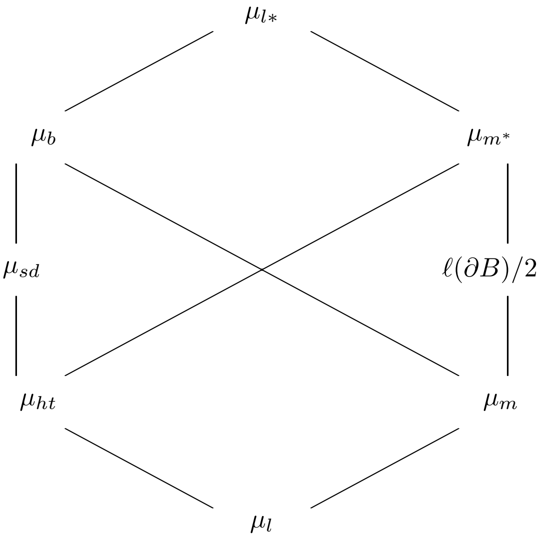
- First of all, there is an order relation between measures on $\mathbb{R}^n$.

$$\mu_1 \leq \mu_2 \iff \mu_1(B) \leq \mu_2(B) \ \forall B.$$

- The set of normalizations is a lattice with this order.

$$(\mu_1 \wedge \mu_2)(B) := \max(\mu_1(B), \mu_2(B)).$$

- The mapping * is an order reversing map.
- This yields (via various theorems and remarks) the following diagram:


μ_{l*}
μ_b
μ_{m^*}
μ_{sd}
$\ell(\partial B)/2$
μ_{ht}
μ_m
μ_l

4. Monotonicity.

We have already seen that:

1. Area should be invariant under isometries.
2. Area should coincide with the usual area when the norm is Euclidean. We now add the following requirement:
3. Area should be monotonic: "smaller lengths gives rise to smaller area".

4. Monotonicity.

To be a bit more precise:
We would like the following to be true.

- If $\mathbb{R}^2$ is given two norms $\|\cdot\|_1$, $\|\cdot\|_2$ coming from unit discs B_1 B_2 and if $B_1 \supseteq B_2$ so that $\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2$ for all vectors $\mathbf{x}$ then we would like the area of a figure in the space $(\mathbb{R}^2, \|\cdot\|_1)$ to be no larger than the area of the identical figure in $(\mathbb{R}^2, \|\cdot\|_2)$.
- One reason for this is aesthetic, it seems reasonable that length should impose some restriction on area.
- A second reason is more technical and comes from the notions of category theory.
  The morphisms are “short” maps.

Consequences of monotonicity

1. A particular normalization is monotonic if and only if its dual is.
2. $\mu_p := \ell(\partial B)/2$ and its dual are not monotonic.
3. If a particular normalization is monotonic then it is continuous from the Banach-Mazur distance.
4. The normalizations μ_I and μ_{I*} are (respectively) the smallest and largest monotonic normalizations.

Consequences of monotonicity

Before coming to the proofs of these statements, there is a simple Proposition.

Proposition A normalization μ is monotonic if and only if

$$\frac{\mu(B_1)}{\mu(B_2)} \geq \frac{\lambda(B_1)}{\lambda(B_2)}$$

for all B_1, B_2 with $B_1 \subseteq B_2$ and λ some fixed Haar measure.

Proof The normalization μ generates measures μ_i on $(\mathbb{R}^n, B_i)$ via the equations

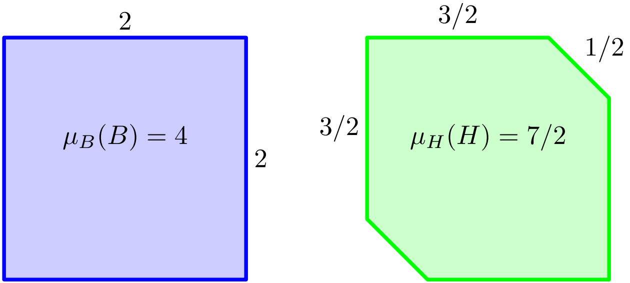
$$\mu_i(A) := \lambda(A)\mu(B_i)/\lambda(B_i), \quad i = 1, 2.$$

Then $\mu_1 \geq \mu_2$ if and only if

$$\frac{\mu(B_1)}{\lambda(B_1)} \geq \frac{\mu(B_2)}{\lambda(B_2)} \iff \frac{\mu(B_1)}{\mu(B_2)} \geq \frac{\lambda(B_1)}{\lambda(B_2)}.$$

Proofs

1. This is easy.
2. Let H be the hexagon with vertices at
   $\pm(1,-1), \ \ \pm(1,1/2) \ \ \pm(1/2,1)$.
   Let B be the usual square with vertices $(\pm 1, \pm 1)$. Then
   $H \subset B$.



Proofs

$\ell(\partial H) = 7$, $\ell(\partial B) = 8$.
Hence we have

$$\frac{\mu(H)}{\mu(B)} = \frac{7}{8} < \frac{15}{16} = \frac{\lambda(H)}{\lambda(B)}.$$

The proposition now shows that μ is not monotonic.

Proofs

3. Let μ be a monotonic normalization.
Suppose B_1 and B_2 are two unit balls which are close in the Banach-Mazur metric, Δ. Then, for some T and small η:

$$B_1 \subseteq T(B_2) \subseteq (1+\eta)B_1.$$

Then, from the Proposition,

$$\frac{\mu(B_1)}{\mu(B_2)} = \frac{\mu(B_1)}{\mu(TB_2)} \geq \frac{\lambda(B_1))}{\lambda(TB_2)} \geq (1+\eta) - n.$$

Also, from the Proposition,

$$\frac{\mu(B_2)}{\mu(B_1)} = \frac{\mu(TB_2)}{\mu((1+\eta)B_1)} \geq \frac{\lambda(T(B_2))}{(1+\eta)^{-n}\lambda(B_1)} \geq (1+\eta)^{-n}.$$

Therefore $|\log(\mu_1(B_1)) - \log(\mu_2(B_2))| < n\log(1+\eta) \leq n\Delta((\mathbb{R}^n, B_1), (\mathbb{R}^n, B_2))$. Hence $|\mu(B_1) - \mu(B_2)| < \epsilon$ for $\Delta((\mathbb{R}^n, B_1), (\mathbb{R}^n, B_2))$ sufficiently small.

4. Let μ be a monotonic normalization and let E be the maximal ellipse inscribed to an arbitrary unit ball B. Then, from the Proposition,

$$\frac{\mu(B)}{\mu(E)} \leq \frac{\lambda(B)}{\lambda(E)}.$$

But $\mu(E) = \pi$ so

$$\mu(B) \leq \pi\lambda(B)/\lambda(E) = \mu_{\ell*}(B).$$

Similarly for the other inequality.

Convexity

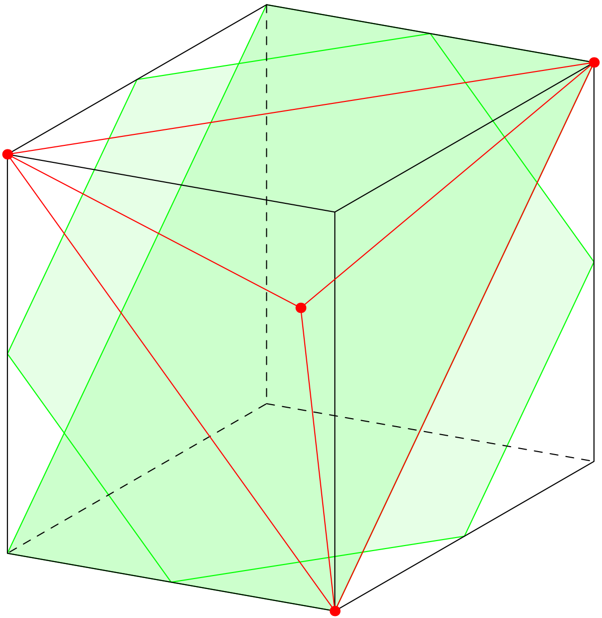
We now add a further requirement:

4. Area should be convex; it should satisfy a simplex inequality.

If S is any simplex in $\mathbb{R}^3$ the area of any one facet should not exceed the sum of the areas of the other three facets.

The normalizations μ_m and μ_l are not convex.

Consider $\mathbb{R}^3$ with the cube B with vertices at $(\pm 1, \pm 1, \pm 1)$ as unit ball. Consider the simplex S with vertices at
$(0, 0, 0)$, $(-1, 1, 1)$, $(1, -1, 1)$ and $(1, 1, -1)$



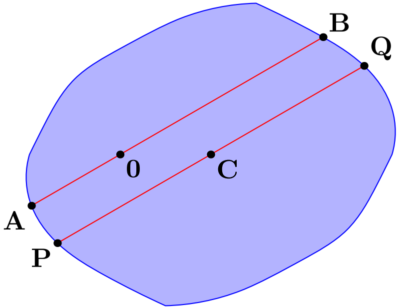
- The four normals to the facets of S are $(1,1,1)$, $(0,1,1)$, $(1,0,1)$ and $(1,1,0)$.
- The cross-sections of B orthogonal to these directions are: (i) a regular hexagon of (Euclidean) side $\sqrt{2}$; and (ii) three rectangles of size $2 \times 2\sqrt{2}$.
- The facet parallel to the hexagon is an equilateral triangle of side $2\sqrt{2}$ and hence area $2/3$ that of the hexagon.
- The other facets are each $1/4$ the area of the cross-section.

1. Using μ_m. The hexagonal cross-section has area 3 and the rectangular cross-sections area 2.
   Hence the facet with normal $(1,1,1)$ has area 2 and the others have area $1/2$.
   μ_m is not convex.
2. Using μ_l. Again, the rectangular cross-sections have area 2 and the corresponding facets area $1/2$.
   The hexagonal cross-section has area $3\sqrt{3}/2$ and the corresponding facet has area $\sqrt{3}$.
   Since $\sqrt{3} > 3/2$, μ_l is also not convex.

5. Non-symmetric norms.

- In this final section let us look at the possibility of norms (unit discs) that are not symmetric.
- For example: a symmetric disc B but with the origin not at the centre.

5. Non-symmetric norms.



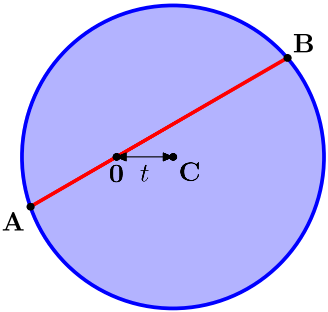
Since the longest chord in each direction is the one through the centre,

$$AO + OB \leq PC + CQ.$$

On average the radial function is smaller than if the origin were at C and so distances are, on average, larger than if the origin were at C.

An extension of the monotonicity requirement would require that, in this case, areas also be larger.

5. Non-symmetric norms.



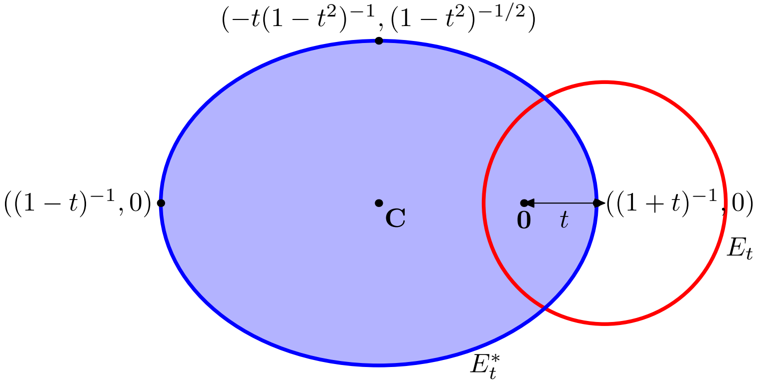
- This can be made particularly precise for the Euclidean disc. Let E_t denote the Euclidean disc with the origin at a distance t from the centre. In this case $OA.OB = (1 - t^2)$; on average (geometric mean) distances increase (compared to the Euclidean case) by a factor of $(1 - t^2)^{-1/2}$.
- The (extension of) monotonicity requires that area also increase in comparison to the Euclidean case.

5. Non-symmetric norms.

- The only definitions which evidently do so are μ_{ht} and μ_{sd}.
- Modify μ_m and μ_{I*} by requiring that the inscribed figures be centred at $\mathbf{0}$ (but μ_m is NOT convex).
- For μ_b a slight modification gives strict inequality.
- Open Question. If K is a non-symmetric convex set, $\mathrm{vp}(K)$ is minimal when the origin is at the Santaló point. Is it true that the mean (in some sense) radial function is maximal (and hence the mean norm is minimal) when the origin is at the Santaló point?

5. Non-symmetric norms.

- The dual of E_t is an ellipse E_t^* with one focus at the origin.



- The semi-major axis is of length $1/(1-t^2)$ and the semi-minor axis is $1/\sqrt{(1-t^2)}$.
- Thus $\mathrm{vp}(E_t) = \pi^2/(1-t^2)^{3/2}$ and $\mu(E_t) = \pi/(1-t^2)^{3/2}$.
- Therefore area increases by a factor of $1/(1-t^2)^{3/2}$ compared to the Euclidean case.
- Thus μ_{ht} looks like a good bet.

5. Non-symmetric norms.

- However, instead of looking at these non-symmetric convex sets let's look at their duals.
- As we just saw, E_t^* is an ellipse with area $\pi/(1-t^2)^{3/2}$ and $\mathrm{vp}(E_t^*) = \pi^2/(1-t^2)^{3/2}$ so that $\mu(E_t^*) = \pi/(1-t^2)^{3/2}$.
- Therefore, μ_{ht} is constant for every t and coincides with the Euclidean measure.
- On the other hand, these unit discs appear to be getting larger and the metrics, therefore, correspondingly smaller which would indicate that mu_{ht} is also not monotonic.
- This is not exactly true.

5. Non-symmetric norms

- let **p** be a point of the dual space with Euclidean norm $\|\mathbf{p}\|_e = 1$.
  Then the norm $\|\mathbf{p}\|^*$ corresponding to E_t^* is the support function of E_t evaluated at **p**. Similarly for $-\mathbf{p}$.
  These are the (Euclidean) distances from the origin to the tangents to E_t determined by **p** and $-\mathbf{p}$.
- Hence, $\|\mathbf{p}\|^* + \|-\mathbf{p}\|^*$ is the (Euclidean) width of E_t in the direction **p** which is constant and equal to 2.
- Therefore, the average of these norms is constantly 1. On average, the norm $\|.\|^*$ is the same as the Euclidean norm and it quite reasonable that the area should also be the same.
- What is needed is a clear definition of what we mean by saying one norm (or unit ball) is larger, on average than another. However, it does look as though μ_{ht} is a good (the best??) normalisation to use for asymmetric norms.

1. Álvarez Paiva, J.C., and Thompson, A.C., Volumes on normed and Finsler spaces, in *A Sampler of Riemann-Finsler Geometry* eds: D. Bao, R.L. Bryant, S-S Chern, and Z. Shen, Volume 50 MSRI Publications, Cambridge, (2004) pp. 1–48.
2. Benson, R. V., *The Geometry of Affine Areas* Ph.D. Thesis, University of Southern California, Los Angeles, (1962). (University Microfilms Inc., Ann Arbor, Michigan 62-6037).
3. Gromov, M., Filling Riemannian manifolds, *J. Diff. Geom.* **18** (1983), 1–147.
4. Holmes, R.D. and Thompson, A.C., N-dimensional area and content in Minkowski spaces, *Pac. J. Math.* **85** (1979) 77–110.
5. Martini, H. and Swanepoel, K. J., Equiframed curves—A generalization of Radon curves, *Monatsh. Math.* **141** (2004) 301–304.
6. McKinney, J.R., On maximal simplices inscribed to a central convex set, *Mathematika* **21** (1974), 38–44.
7. Mustafaev, Z., The ratio of the length of the unit circle to the area of the unit disc on Minkowski planes, *Proc. Amer. Math. Soc.* **133** (2005) 1231–1237.
8. Petty, C.M., On the geometry of the Minkowski plane, *Riv. Mat. Univ. Parma* (4) **6** (1955), 269–292.