Symmetrization and the Mazur-Ulam theorem

Matias Dahl

February 7, 2007

Abstract

The Mazur-Ulam theorem states that a bijective isometry between two normed spaces is an affine map. This note is based on the observation that symmetrization of norms preserve isometries. Using this observation, we show that the Mazur-Ulam theorem for non-symmetric norms follow directly from the Mazur-Ulam theorem for symmetric norms (Section 1). In addition, using this observation we present a direct proof of the Mazur-Ulam theorem for smooth non-symmetric Minkowski norms (Section 2). The notion of Minkowski norms in this note bears no connection with non-definite inner products in relativity.

1 Normed spaces

Suppose V is a real vector space. By a *norm* on V we mean a function $F\colon V \to \mathbb{R}$ such that

1. $F(v) \geq 0$ for all $v \in V$ and $F(v) = 0$ if and only if $v = 0$.
2. $F(\lambda v) = \lambda F(v)$ for all $v \in V$ and $\lambda > 0$.
3. $F(u + v) \leq F(u) + F(v)$ for all $u, v \in V$.

If $F(v) = F(-v)$ for all $v \in V$, then F is *symmetric*. If (V, F) and (W, G) are two normed spaces and $\Psi\colon V \to W$ is a map, then Ψ is an *isometry* if

$$G(\Psi(a) - \Psi(b)) = F(a - b), \quad a, b \in V. \tag{1}$$

Also, a map Ψ between vector spaces is *affine* if $\Psi - \Psi(0)$ is linear.

Example 1.1. Suppose (V, F) is a normed space, and $\Psi\colon V \to W$ is a bijective affine map. Then Ψ is an bijective isometry between normed spaces (V, F) and $(W, F \circ (\Psi - \Psi(0))^{-1})$.

The Mazur-Ulam theorem shows the converse of the above example:

Theorem 1.2 (Mazur-Ulam, 1932). *A bijective isometry between two normed spaces is affine.*

Let us point out that if an isometry is a surjection, it is a bijection. Therefore the Mazur-Ulam theorem is sometimes also formulated for surjective isometries. Proofs for the Mazur-Ulam theorem for symmetric norms can be found in [Lax02, Tho96, V03̈]. Our next aim is to show that this symmetric Mazur-Ulam theorem implies the non-symmetric Mazur-Ulam theorem, that is, Theorem 1.2. For this purpose, let us define the *symmetrization* of a norm $F\colon V \to \mathbb{R}$ as

$$\hat{F}(v) = \frac{F(v) + F(-v)}{2}, \quad v \in V.$$

It is not difficult to check that $\hat{F}$ is always a symmetric norm. A direct calculation also gives the following result:

Proposition 1.3. *Suppose (V, F) and (W, G) are normed spaces, and $\hat{F}$ and $\hat{G}$ are symmetrizations of F and G, respectively. If $\Psi\colon (V, F) \to (W, G)$ is an isometry, then $\Psi\colon (V, \hat{F}) \to (W, \hat{G})$ is an isometry.*

Proposition 1.3 now reduces the non-symmetric Mazur-Ulam theorem to the symmetric Mazur-Ulam theorem; If $\Psi\colon (V, F) \to (W, G)$ is a bijective isometry between normed spaces, then $\Psi\colon (V, \hat{F}) \to (W, \hat{G})$ is a bijective isometry between symmetric normed spaces, and by the symmetric Mazur-Ulam theorem, Ψ is affine.

2 Minkowski spaces

Hereafter we only work with finite dimensional vector spaces. We shall never explicitly write out the basis. However, if v is a vector, we denote by v^i its components in some unspecified fixed basis. Also, $\frac{\partial f}{\partial v^i}$ denotes partial differentiation with respect to these basis vectors. The Einstein summing convention is used throughout.

Let V be a finite dimensional real vector space. Then a *Minkowski norm* on V is a function $F\colon V \to [0, \infty)$ such that [BCS00, She01].

1. F is smooth on $V \setminus \{0\}$.
2. F is positively 1-homogeneous; for $v \in V$ and $\lambda > 0$,

$$F(\lambda v) = \lambda F(v).$$

3. F is strongly convex; for each $v \in V \setminus \{0\}$,

$$g_{ij}(v) = \frac{1}{2} \frac{\partial^2 F^2}{\partial v^i \, \partial v^j}(v)$$

is a positive definite matrix.

If (V, F) and (W, G) are Minkowski spaces, and $\Psi \colon V \to W$ is smooth map, then Ψ is an *isometry* provided that equation (1) holds. We can now state the Mazur-Ulam theorem for Minkowski norms:

Theorem 2.1 (Mazur-Ulam, 1932). *A bijective isometry between two Minkowski spaces is affine.*

One can prove that every Minkowski norm satisfies the triangle inequality, and that every Minkowski norm is a norm in the above sense [BCS00, She01]. Hence Theorem 2.1 is a corollary of Theorem 1.2. Our next aim is to give a direct proof of Theorem 2.1. For this, we need some preliminaries.

Suppose r is an integer, and $f \colon V \to \mathbb{R}$ is smooth on $V \setminus \{0\}$. Then f is *positively r-homogeneous* if

$$f(\lambda v) = \lambda^r f(v), \quad v \in V, \quad \lambda > 0.$$

Similarly, f is *absolutely r-homogeneous* if

$$f(\lambda v) = |\lambda|^r f(v), \quad v \in V, \quad \lambda \in \mathbb{R} \setminus \{0\}.$$

An important property of Minkowski norms is that each $y \in V \setminus \{0\}$ induces an inner product $g_y(u, v) = g_{ij}(y) u^i v^j$. Symmetric Minkowski norms are defined as for norms. However, we shall define the symmetrization of Minkowski norms slightly differently than for norms. Instead of symmetrizing the norm, we symmetrize the inner product.

Proposition 2.2 (Symmetrization of Minkowski norms). *Suppose (V, F) is a Minkowksi space. Then*

$$\hat{F}(v) = \sqrt{\frac{F^2(v) + F^2(-v)}{2}}, \quad v \in V$$

is a symmetric Minkowski norm for V. (Here, $\sqrt{\cdot}$ is the positive square root.)

Proof. This follows as the sum of two positive matrices is again positive definite. □

For Minkowski norms, the analogue of Proposition 1.3 is:

Proposition 2.3. *Suppose (V, F) and (W, G) are Minkowski spaces, and $\hat{F}$ and $\hat{G}$ are symmetrizations of F and G, respectively. If $\Psi\colon (V, F) \to (W, G)$ is an isometry, then $\Psi\colon (V, \hat{F}) \to (W, \hat{G})$ is an isometry.*

If F is a symmetric Minkowski norm, then F^2 is absolutely 2-homogeneous, and the following relations hold for all $v \in V \setminus \{0\}$ and $\lambda \neq 0$,

$$\frac{\partial F^2}{\partial v^i}(\lambda v) = \lambda \frac{\partial F^2}{\partial v^i}(v), \tag{2}$$

$$\frac{\partial^2 F^2}{\partial v^i \partial v^j}(\lambda v) = \frac{\partial^2 F^2}{\partial v^i \partial v^j}(v), \tag{3}$$

$$\frac{\partial^2 F^2}{\partial v^i \partial v^j}(v) v^j = \frac{\partial F^2}{\partial v^i}(v). \tag{4}$$

Equations (2)-(3) follow by differentiating $F^2(\lambda v) = \lambda^2 F^2(v)$ with respect to v^i. Equation (4) is known as *Euler's theorem* and it follows by differentiating equation (2) with respect to λ.

Proof of Theorem 2.1. Let $\Psi\colon (V, F) \to (W, G)$ be the isometry, let $a, b \in V \setminus \{0\}$, $a \neq b$ be arbitrary, and let $\Delta_{uv} = \Psi(u) - \Psi(v)$ for $u, v \in V$. By Proposition 2.3, we may assume that F and G are symmetric. Let v^i and w^i be coordinates for V and W, respectively. Differentiating the square of equation (1) with respect to a^i gives

$$\begin{aligned} &\frac{\partial^2 F^2}{\partial v^i \partial v^j}(a - b) \\ &\quad = \frac{\partial^2 G^2}{\partial w^l \partial w^m}(\Delta_{ab}) \frac{\partial \Psi^l}{\partial v^i}(a) \frac{\partial \Psi^m}{\partial v^j}(a) + \frac{\partial G^2}{\partial w^l}(\Delta_{ab}) \frac{\partial^2 \Psi^l}{\partial v^i \partial v^j}(a). \end{aligned}$$

Let us fix $a, c \in V \setminus \{0\}$, and let $b = b(t) \in V$ be the solution to $\Delta_{ab(t)} = tc$ for $t \in \mathbb{R} \setminus \{0\}$. Then $a \neq b(t)$, so the right hand side of the above equation is always an invertible matrix. Let A be the first term. It is positive definite, and by equation (3), A is independent of t. Similarly, by equation (2), the second term depends linearly on t. Hence it equals tB for a symmetric matrix B, and $A + tB$ is invertible for $t \neq 0$. By Lemma 2.4, $B = 0$. Thus, by equation (4), the Hessians of Ψ^l must vanish for all l and $a \in V \setminus \{0\}$. As Ψ is smooth, the Hessians must also vanish for $a = 0$. The result follows from Lemma 2.5. ☐

Lemma 2.4. *Suppose A, B are real symmetric matrices, and A is positive definite. If $A + tB$ is invertible for all $t \in \mathbb{R} \setminus \{0\}$, then $B = 0$.*

Proof. Let us first show that $\sigma(A^{-1}B)$, the *spectrum* of $A^{-1}B$, is real. To see this, let S be the positive definite square root of A. Then

$$\sigma(A^{-1}B) = \sigma(SA^{-1}BS^{-1}) = \sigma(S^{-1}BS^{-1}), \tag{5}$$

and the claim follows since S^{-1} is symmetric. Next, by the assumption on $A + tB$, $A^{-1}B - tI$ is invertible for all $t \neq 0$. Thus $\sigma(A^{-1}B) \subseteq \{0\} \cup \mathbb{C} \setminus \mathbb{R}$. Thus $\sigma(A^{-1}B) = 0$, and by equation (5), $B = 0$ as $S^{-1}BS^{-1}$ is diagonalizable. ☐

Lemma 2.5. *A smooth function $f\colon \mathbb{R}^n \to \mathbb{R}$ is affine if and only if its Hessian is identically zero.*

Proof. If the gradient of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ vanishes identically, then f is constant. Thus, if $x \in \mathbb{R}^n$, and γ is a smooth curve from 0 to x, then $0 = \int_\gamma df = f(x) - f(0)$. See [Con96], p. 168. By assumption,

$$\frac{\partial}{\partial x^i}\left(\frac{\partial f}{\partial x^j}\right)(x) = 0$$

for all $i, j = 1, \ldots, n$ and $x \in \mathbb{R}^n$, so

$$\frac{\partial f}{\partial x^j}(x) = C_j$$

for some constants $C_1, \ldots, C_n$. Thus

$$\frac{\partial}{\partial x^j}\left(f(x) - \sum_{l=1}^n C_l x^l\right) = 0$$

and the result follows. ☐

References

[BCS00] D. Bao, S.-S. Chern, and Z. Shen, *An introduction to Riemann-Finsler geometry*, Springer, 2000.

[Con96] L. Conlon, *Differentiable manifolds, A first course*, Birkhäuser, 1996.

[Lax02] P. D. Lax, *Functional analysis*, Wiley-Interscience, 2002.

[She01] Z. Shen, *Lectures on Finsler geometry*, World Scientific, 2001.

[Tho96] A. C. Thompson, *Minkowski geometry*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1996.

[V03] J. Väisälä, *A proof of the Mazur-Ulam theorem*, The American Mathematical Monthly **110** (2003), no. 7, 633–635.