Norms, Isometries, and Isometry Groups

Chi-Kwong Li

1 Introduction

The study of linear algebra has become more and more popular in the last few decades. People are attracted to this subject because of its beauty and its connections with many other pure and applied areas. In theoretical development of the subject as well as in many applications, one often needs to measure the "length" of vectors. For this purpose, norm functions are considered on a vector space. In this expository article, we explain why one would want to study different kinds of norms on a real vector space. We then focus on the problem of how to identify different norms using linear isomorphisms with particular attention to a group theory method used by several authors recently.

One may view the first half of this article as a gentle introduction to the theory of norms and the second half as an illustration of how group theory can be applied to questions in linear algebra.

2 What are norms and why study them?

A *norm* on a real vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ satisfying:

$\|u\| > 0$ for any nonzero $u \in V$, (positive)

$\|ru\| = |r|\|u\|$ for any $r \in \mathbb{R}$ and $u \in V$, (homogeneous)

$\|u + v\| \le \|u\| + \|v\|$ for any $u, v \in V$. (triangle inequality)

The homogeneous condition ensures that the norm of the zero vector in V is 0; this condition is often included in the definition of a norm.

Common examples of norms on $\mathbb{R}^n$ are the ℓ_p *norms*, where $1 \le p \le \infty$, defined by

$$\ell_p(x) = \left\{ \sum_{j=1}^{n} |x_j|^p \right\}^{1/p} \qquad \text{if} \quad 1 \le p < \infty, \text{ and} \tag{1}$$

$$\ell_p(x) = \max_{1 \le j \le n} |x_j| \qquad \text{if} \quad p = \infty,$$

for any $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$. Note that if one defines an ℓ_p function on $\mathbb{R}^n$ as in (1) with $0 < p < 1$, then it does not satisfy the triangle inequality, and hence is not a norm.

Other examples of norms on $\mathbb{R}^n$ are the *c-norms* defined by

$$\|x\|_c = \max\{x^t P c : P \in GP(n)\},$$

in which $c \in \mathbb{R}^n$ is nonzero and $GP(n)$ denotes the set of all *generalized permutation matrices* (also known as *signed permutation matrices* or *monomial matrices*), the matrices with exactly one nonzero entry equal to ± 1 in each row and column.

Norms can be regarded as generalizations of the absolute value function of numbers. Actually, one easily verifies:

Fact 1 *Consider* $\mathbb{R}$ *as a real vector space. The absolute value function on* $\mathbb{R}$ *is a norm, and every norm on* $\mathbb{R}$ *is a positive scalar multiple of the absolute value.*

With the absolute value function on $\mathbb{R}$, one can compare the magnitudes of numbers, discuss the convergence of sequences, study limits and continuity of functions, and consider approximation problems such as finding the nearest integer or prime to a given real number. The same is true for a norm on a vector space. Given a norm on a real vector space V, one can compare the norms of vectors, discuss convergence of sequences of vectors, study limits and continuity of transformations, and consider approximation problems such as finding the nearest element in a subset or a subspace of V to a given vector. These problems arise naturally in analysis, Lie theory, numerical analysis, differential equations, Markov chains, econometrics, population models in biology or sociology, equilibrium states in physics and chemistry; see [1], [9], [10], [18] and their references.

As pointed out in Fact 1, there is only one norm on $\mathbb{R}$ up to positive scalar multiples. In general, one may have a much wider variety of norms on a vector space. For example, we have the ℓ_p norms on $\mathbb{R}^n$, which are not scalar multiples of each other when $n \geq 2$. Actually, it is not hard to verify:

Fact 2 *All norms on a vector space* V *are positive scalar multiples of a single given norm if and only if* V *is one dimensional.*

3 Why study different norms?

Different norms on a vector space can give rise to different geometrical and analytical structures. In an infinite dimensional vector space, the convergence of a sequence can vary, depending on the choice of norm. This phenomenon leads to many interesting questions and research in analysis and functional analysis; see [16] and [17].

In a finite dimensional vector space V, all norms are *equivalent* in the following sense: for any norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on V, there are positive constants a and b such that

$$a\|v\|_1 \leq \|v\|_2 \leq b\|v\|_1 \quad \text{for all } v \in V.$$

(A fancy description is that there is only one norm topology on a finite dimensional vector space.) Consequently, the convergence of a sequence of vectors in a finite dimensional vector space is independent of the choice of norm. Nevertheless, there are reasons to consider different norms.

First, for a given sequence it may be easier to prove convergence with respect to one norm rather than another. In applications such as numerical analysis, one would like to use a norm that can determine convergence efficiently. Therefore, it is a good idea to have knowledge of different norms.

Second, sometimes a specific norm may be needed to deal with a certain problem. For instance, if one travels in Manhattan and wants to measure the distance from a location marked as the origin $(0, 0)$ to a destination marked as (x, y) on the map, one may use the

ℓ_2 norm of (x, y), which measures the straight line distance between the two points, or one may need to use the ℓ_1 norm of v, which measures the distance for a taxi cab to drive from $(0, 0)$ to (x, y). The ℓ_1 norm is sometimes referred to as the *taxi cab norm* for this reason.

In approximation theory, solutions of a problem can vary with different norms. For example, if W is a subspace of $\mathbb{R}^n$ and $v \notin W$, then for $1 < p < \infty$ there is a unique $u_0 \in W$ such that

$$\|v - u_0\| \leq \|v - u\| \qquad \text{for all } u \in W,$$

but the uniqueness condition may fail if $p = 1$ or ∞. To see a concrete example, let $v = (1, 0)$ and $W = \{(0, y) : y \in \mathbb{R}\}$. Then for all $y \in [-1, 1]$ we have $1 = \|v - (0, y)\| \leq \|v - w\|$ for all $w \in W$. For some problems, having a unique approximation is good, but for others it may be better to have many so that one of them can be chosen to satisfy additional conditions.

4 Identify two norms

While we emphasis the importance of studying different norms, one should avoid wasting effort in studying two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on V that are essentially the same in the sense that there is a linear bijection $L : V \to V$ satisfying

$$\|L(v)\|_2 = \|v\|_1 \qquad \text{for all } v \in V. \tag{2}$$

More generally, one can identify two normed vector spaces $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$ if there is a linear bijection $L : V_1 \to V_2$ so that (2) holds for all $v \in V_1$. We call such an L an *isometric isomorphism* between $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$. If $V_1 = V_2$ and $\|\cdot\|_1 = \|\cdot\|_2$, we say that such an L is an *isometry* for $\|\cdot\|_1$ on V_1.

Studying bijections that preserve the basic structure of a mathematical system always leads to better understanding of the system. This is a reason why one would study linear isomorphisms between vector spaces, homeomorphisms between topological spaces, and so forth. We illustrate this general comment in the context of normed vector spaces in the following.

Since the ℓ_2 norm has a lot of symmetries, there are a lot of isometries, namely, all the orthogonal matrices, for it. If $p \neq 2$, there are not so many symmetries for the ℓ_p norm and only generalized permutation matrices can be isometries for it. For simple proofs of these facts, see [4], [12], [15].

Many techniques have been developed to characterize isometries for a given norm. These methods often involve knowledge from other areas [7]. In the forthcoming sections, we discuss a group theory method, which was not mentioned in [7], and some related results.

5 Study isometries using group theory

A basic result on isometries asserts that the collection of isometries for a given norm is a group, a subgroup of the group of invertible operators under composition of functions. To verify this, it suffices to check that the identity map is always an isometry, and if L_1 and L_2 are isometries of a norm then so is $L_1^{-1}L_2$.

Certainly, knowing the isometries gives complete information about the isometry group. For example, by the previous discussion we see that the group $O(n)$ of orthogonal matrices is the isometry group of the ℓ_2 norm, and if $p \neq 2$ then the group $GP(n)$ of generalized permutation matrices is the isometry group of the ℓ_p norm.

In the following, we show that sometimes it is easy to characterize the isometry group directly; the result can then be used to determine the structure of isometries. This, in a certain sense, illustrates a principle of Polya [14]: *If you wish to prove a theorem, it is sometimes easier to prove a harder theorem* (*that covers the original theorem as a special case*)*!*

Using the group theory approach, one may borrow mathematical tools from other subjects such as the theory of Lie groups and algebraic groups. Moreover, getting complete information about isometry groups may help solve other problems involving the given norms; see [5], [8], [13]. We describe some results in the following to illustrate our point.

A *symmetric norm* (also known as a *symmetric gauge function*) is a norm on $\mathbb{R}^n$ such that

$$\|Px\| = \|x\| \quad \text{for all } P \in GP(n) \text{ and all } x \in \mathbb{R}^n.$$

Evidently, ℓ_p norms and c-norms are symmetric norms. Suppose one focuses on the isometry group G of a symmetric norm. It is easy to see that G must be closed and bounded, i.e., compact, in $\mathbb{R}^{n\times n}$, and it contains $GP(n)$ as a subgroup. This leads naturally to the question of determining the compact groups of $n \times n$ real matrices that contains $GP(n)$. To this end we have

Theorem 1 *A compact group of matrices in $\mathbb{R}^{n\times n}$ that contains $GP(n)$ must be one of the following:*

(a) $O(n)$,

(b) $GP(n)$,

(c) *if $n = 2$, the dihedral group D_{8k} with $8k$ elements for some integer $k \geq 1$, or*

(d) *if $n = 4$, a group generated by $GP(4)$ and one of the following matrices*

$$A = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}, \quad B = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}. \tag{3}$$

Consequently, one sees that the isometry group of a symmetric norm must be one of the groups described in Theorem 1. Of course, only a multiple of the ℓ_2 norm can have $O(n)$ as the isometry group. For other symmetric norms, the isometry group must be $GP(n)$ unless $n = 2$ or 4. In these exceptional cases, one can decide easily which one should be the isometry group for a given symmetric norm.

It is worth mentioning that Theorem 1 can be deduced from the theory of reflection groups and Lie groups ([2], [3]), but it can also be proved by elementary methods [6].

6 Isometric isomorphisms

Knowledge of isometry groups can help us determine efficiently the isometric isomorphisms relating two normed spaces. For such problems, the following observation is useful.

Observation A *Suppose G_j is the isometry group of $\|\cdot\|_j$ on V_j for $j = 1, 2$. If $L : V_1 \to V_2$ is an isometric isomorphism, then $L^{-1}G_2L = G_1$.*

If $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$ are isometrically isomorphic, there is no loss of generality to assume that $V_1 = V_2$. Theorem 1 gives complete information about the possible isometry groups for symmetric norms on $\mathbb{R}^n$. In particular, the several types of groups identified in Theorem 1 have different cardinality. Thus, if G_1 and G_2 are isometry groups of symmetric norms on $\mathbb{R}^n$ and there exists an invertible operator L such that $L^{-1}G_2L = G_1$, then G_2 and G_1 must be the same. Hence, the problem reduces to studying the existence and the structure of an L such that $L^{-1}GL = G$ for those groups G in Theorem 1. Recall that if G_1 is a subgroup of G_2, then the *normalizer* of G_1 in G_2 is the collection of elements $g \in G_2$ such that $g^{-1}G_1g = G_1$. Hence, our problem is to find the normalizer of G in the group of invertible operators on $\mathbb{R}^{n\times n}$. With some effort, one can prove the following.

Theorem 2 *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be symmetric norms on $\mathbb{R}^n$. Then L is an isometric isomorphism between $(\mathbb{R}^n, \|\cdot\|_1)$ and $(\mathbb{R}^n, \|\cdot\|_2)$ if and only if there is a $\gamma > 0$ such that one of the following holds:*

(a) *$G_1 = G_2$, $\|x\|_1 = \gamma\|x\|_2$ for all $x \in \mathbb{R}^n$, and $\gamma^{-1}L \in G_1$.*

(b) *$n = 4$, $G_1 = G_2 = \langle GP(4), A\rangle$, the group generated by $GP(4)$ and A, $\|x\|_1 = \gamma\|x\|_2$ for all $x \in \mathbb{R}^4$, and $\gamma^{-1}L \in \langle GP(4), B\rangle \setminus \langle GP(4), A\rangle$, where A and B are the matrices in (3).*

(c) *$n = 2$, $G_1 = G_2 = D_{8k}$, $\|x\| = \gamma\|Rx\|$ for some $R \in D_{16k} \setminus D_{8k}$, and $\gamma^{-1}L \in D_{16k} \setminus D_{8k}$.*

Theorem 2 illustrates how difficult it is for $\mathbb{R}^n$ with two different symmetric norms to be isometrically isomorphic.

7 Identify a norm with its dual

Suppose V is a real inner product space with inner product $(\cdot,\cdot)$, and suppose $\|\cdot\|$ is a given norm on V, which need not be the norm derived from the inner product [9, p. 262]. The dual norm $\|\cdot\|^D$ of $\|\cdot\|$ is defined by

$$\|x\|^D = \max\{|(x, y)| : \|y\| \le 1\}.$$

For example, with the usual Euclidean inner product $(x, y) = y^tx$ on $\mathbb{R}^n$, the ℓ_1 and ℓ_∞ norms are dual to each other, and for $1 < p < \infty$ the dual norm of ℓ_p is the ℓ_q norm, where $p^{-1} + q^{-1} = 1$. Clearly, the ℓ_2 norm is self-dual. More generally, it is known that

$\|x\| = \gamma\|x\|^D$ if and only if $\|x\| = \sqrt{\gamma}\ell_2(x)$ [9, Theorem 5.4.16]. Again, knowledge of the isometry group of a norm is useful in studying the isometric isomorphisms between the norm and its dual norm.

Given a linear operator L on the real inner product space V, the *dual transformation* L^* is the unique linear operator on V satisfying $(x, L^*(y)) = (L(x), y)$ for all $x, y \in V$. If $V = \mathbb{R}^n$ is equipped with the usual inner product, then the dual transformation of $A \in \mathbb{R}^{n\times n}$ is just A^t, the transpose of A. We have the following observation.

Observation B *Suppose G is the isometry group of the norm $\|\cdot\|$ on the inner product space V.*

(a) *The isometry group G^* of $\|\cdot\|^D$ is $\{L^* : L \in G\}$.*

(b) *If L is an isometric isomorphism between $(V, \|\cdot\|)$ and $(V, \|\cdot\|^D)$, then $L^{-1}G^*L = G$.*

Theorem 1 ensures that the isometry group G of a symmetric norm on $\mathbb{R}^n$ always satisfies $G^* = G$. Thus, checking condition (b) in Observation B again reduces to studying an invertible operator L such that $L^{-1}GL = G$. One has the following result.

Theorem 3 *Let $\|\cdot\|$ be a symmetric norm on $\mathbb{R}^n$. Then $(\mathbb{R}^n, \|\cdot\|)$ is isometrically isomorphic to $(\mathbb{R}^n, \|\cdot\|^D)$ if and only if there exists a $\gamma > 0$ such that one of the following holds:*

(a) *$\|x\| = \gamma\ell_2(x)$ for all $x \in \mathbb{R}^n$.*

(b) *$n = 4$, the isometry group of the norm is $\langle G, A\rangle$, and the norm satisfies $\|x\| = \gamma\|Bx\|^D$ for all $x \in \mathbb{R}^4$, where A and B are the matrices in (3).*

(c) *$n = 2$, the isometry group of the norm is D_{8k}, and there exists $R \in D_{16k} \setminus D_{8k}$ such that $\|x\| = \gamma\|Rx\|$ for all $x \in \mathbb{R}^2$.*

8 Retain desired norm properties

Given a norm $\|\cdot\|$ and an invertible linear operator S on V, a standard technique to obtain a new norm is to define

$$\|x\|_S = \|S(x)\| \quad \text{for all } x \in V.$$

If the norm $\|\cdot\|$ has some nice properties, one would like to see that $\|\cdot\|_S$ retains them. In particular, if one views $S(x)$ as a change of basis operation, then one would like the nice norm properties to be preserved by the change of basis.

One way to impose nice properties on a norm is to require it be H-invariant for a certain (compact) group H of linear operators on V [11], i.e.,

$$\|L(x)\| = \|x\| \quad \text{for all } L \in H \text{ and all } x \in V.$$

One would like to characterize those S such that $\|\cdot\|_S$ is also an $H-$invariant norm. We have the following observation.

Observation C *Suppose G is the isometry group of an H-invariant norm $\|\cdot\|$ on V. Then*

(a) $S^{-1}GS$ *is the isometry group of* $\|\cdot\|_S$,

(b) $\|\cdot\|$ *is* H*-invariant if and only if* $H < S^{-1}GS$, *and*

(c) *if* S *satisfies* $S^{-1}GS = G$ *or* $S^{-1}HS = H$, *then* $\|\cdot\|_S$ *is* $H-$*invariant.*

Clearly, symmetric norms are $H-$invariant norms with $H = GP(n)$. Using Observation C, one has:

Theorem 4 *Let* $\|\cdot\|$ *be a symmetric norm on* $\mathbb{R}^n$*, and let* S *be an invertible operator on* $\mathbb{R}^n$*. The norm* $\|\cdot\|_S$ *is a symmetric norm if and only if there exists a* $\gamma > 0$ *such that one of the following holds:*

(a) $\gamma S \in G$,

(b) $n = 4$, $G = \langle GP(4), A\rangle$, *and* $\gamma S = PBQ$ *for some* $P, Q \in GP(4)$, *where* A *and* B *are the matrices in (3), or*

(c) $n = 2$, $G = D_{8k}$ *and* $\gamma S \in D_{16k} \setminus D_{8k}$.

9 Conclusion

We have illustrated the importance of studying different norms on a vector space and have shown how studying the isometry group can help deduce information on isometries for a norm. In many situations, this approach is efficient and allows one to use mathematical tools from other areas. Moreover, using knowledge of isometry groups and some general observations, one can solve several other problems effectively. Related results can be found in [5], [8], [11], [12], [13]. There is much potential for further development and applications of the group theory approach.

References

[1] G.R. Belitskii and Yu. I. Lyubich, *Matrix Norms and Their Applications,* Oper. Theory Adv. Appl. 36, Birkhäuser, Boston, 1988.

[2] C.T. Benson and L.C. Grove, *Finite Reflection Groups,* Springer-Verlag, New York, 1985.

[3] Bourbaki, *Groupes et. Algebras de Lie,* Chap. 4-6, Hermann, Paris, 1968.

[4] M.S. Braverman and E.M. Semenov, Isometries of symmetric spaces, *Soviet Math. Dokl.* 15 (1974) 1027-1030.

[5] D.Z. Djokovic and C.K. Li, Overgroups of some classical linear groups with applications to linear preserver problems, *Linear Algebra Appl.* 197/198 (1994) 31-62.

[6] D.Z. Djokovic, C.K. Li, and L. Rodman, Isometries of symmetric gauge functions, *Linear and Multilinear Algebra* 30 (1991) 81-92.

[7] R.J. Fleming and J.E. Jamison, Isometries on Banach spaces: a survey, pp. 52–123 in *Analysis, geometry and groups: a Riemann legacy volume,* Hadronic Press Collect. Orig. Artic., Hadronic Press, Palm Harbor, FL, 1993.

[8] R.M. Guralnick and C.K. Li, Invertible preservers and algebraic groups III: Preservers of unitary similarity (congruence) invariants and overgroups of some unitary subgroups, *Linear and Multilinear algebra* 43 (1997) 257-282.

[9] R.A. Horn and C.R. Johnson, *Matrix Analysis,* Cambridge University Press, New York, 1985.

[10] A.S. Householder, *The Theory of Matrices in Numerical Analysis,* Blaisdell, New York, 1964.

[11] C.K. Li, Some aspects of the theory of norms, *Linear Algebra Appl.* 212/213 (1994) 71-100.

[12] C.K. Li and W. So, Isometric isomorphisms between normed spaces, *Rocky Mountain J. Math.* 28 (1998) 607-624.

[13] C.K.Li and W. Whitney, Closed symmetric overgroups of S_n in O_n, *Canad. Math. Bull.* 39 (1996) 83-94.

[14] G. Polya, *How to Solve It,* Princeton University Press, Princeton, 1971.

[15] S. Rolewicz, *Metric Linear Spaces,* 2nd ed., PWN-Polish Scientific Publications, Warsaw, 1985.

[16] W. Rudin, *Real and complex analysis,* 3rd ed., McGraw-Hill, New York, 1987.

[17] W. Rudin, *Functional analysis,* 2nd ed., McGraw-Hill, New York, 1991.

[18] I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces,* Springer-Verlag, New York, 1970.

College of William and Mary, P.O. Box 8795, Williamsburg, VA 23187-8795, USA
ckli@math.wm.edu