Angles between Euclidean subspaces

Sheng Jiang

Angles Between Euclidean Subspaces

SHENG JIANG
Department of Mathematics, Yangzhou University, Yangzhou, Jiangsu 225002, P.R. China.

Abstract. The angle between two subspaces of dimensions p and q in a Euclidean space is considered by using exterior algebra. Some properties of angles are obtained. The relation between such a higher dimensional angle and the usual principal angles is also given. And finally, an application to Grassmann manifolds is briefly discussed.

Mathematics Subject Classifications (1991): 51M05, 51M16, 51K05.

Key words: higher-dimensional angle; principal angles; Grassmann manifolds.

The trigonometry in a higher-dimensional Euclidean space has been studied by several authors. Döband [1] proved the generalized law of cosines, the Pythagorean theorem and the Heron formula for n-simplices in a Euclidean space by means of determinants. Eriksson [2] generalized the law of sines. S.-Y. T. Lin and Y.-F. Lin [3] gave a new approach of the n-dimensional Pythagorean theorem. And, Miao and Ben-Israel [4] established some new properties for the principal angles between subspaces in a Euclidean n-space. These principal angles have an important application in statistics to the canonical correlation theory of Hotelling [5].

In this paper we will consider the angle between two subspaces in an n-dimensional Euclidean space E^n, derive some new properties, find the relation between the higher dimensional angle and the principal angles of two subspaces, and briefly discuss an example of application to the geometry of Grassmann manifolds.

1. Preliminaries

For convenience, we will use exterior algebra for computations, it makes the results clear and the proofs simpler. In this section we briefly state some basic facts about exterior or Grassmann algebra which are needed in our paper; for details see, for example, Bourbaki [6, ch. 3], or Flanders [7, ch. 2].

Let E^n be an n-dimensional real linear space endowed with a Euclidean inner product. For two vectors a and b, denote by $a \cdot b$ their inner product.

The symbol $\wedge^p E^n$ denotes the linear space consisting of all linear combinations with real coefficients of wedges of p vectors in E^n, $1 \leq p \leq n$. Elements of $\wedge^p E^n$

are called p-vectors over E^n. A p-vector is said to be simple or decomposable, if it can be decomposed as a single wedge of p vectors in E^n.

An induced product in $\wedge^p E^n$ can be naturally introduced as follows.

For a pair of p-vectors

$$\alpha = a_1 \wedge a_2 \wedge \cdots \wedge a_p, \quad \beta = b_1 \wedge b_2 \wedge \cdots \wedge b_p,$$

define

$$\langle \alpha, \beta \rangle = \det (a_i \cdot b_j).$$

We can prove the following

LEMMA 1. *The linear space $\wedge^p E^n$ endowed with $\langle \, , \rangle$ is a Euclidean space.*

Proof. Choosing in E^n an orthonormal basis $\{e_j\}$, $j = 1, 2, \ldots, n$, we get in $\wedge^p E^n$ an induced orthonormal basis $\{E_K\}$, $K = 1, 2, \ldots, \binom{n}{p}$, where the E_K's are the wedges of p different vectors e_j (see Flanders [7, p. 14]). An arbitrary p-vector α can be decomposed as

$$\alpha = \sum_K \lambda_K E_K,$$

where the λ_K's are real scalars. So

$$\langle \alpha, \alpha \rangle = \sum_K (\lambda_K)^2.$$

Hence for any $\alpha \in \wedge^p E^n$, $\langle \alpha, \alpha \rangle = 0$ if and only if $\alpha = 0$. Consequently $\langle \, , \rangle$ defines a Euclidean inner product. □

COROLLARY. *For arbitrary α, $\beta \in \wedge^p E^n$, we have*

$$\langle \alpha, \alpha \rangle \langle \beta, \beta \rangle \geq \langle \alpha, \beta \rangle^2,$$

where equality holds if and only if α differs from β only in a scalar factor.

Remark. The length of a p-vector α is defined by

$$|\alpha| = \sqrt{\langle \alpha, \alpha \rangle}.$$

$\alpha = 0$ if and only if $|\alpha| = 0$.

2. Higher Dimensional Angle

Let $\alpha = a_1 \wedge a_2 \wedge \cdots \wedge a_p$ be a nonzero decomposable p-vector. Then α corresponds to a p-dimensional subspace A^p of E^n spanned by $a_1, a_2, \ldots, a_p$. Conversely, for any basis of the vector subspace A^p, the wedge of basis vectors equals $k\alpha$, where k is a nonzero real number.

Further, suppose that $1 \leq p \leq q < n$, and

$$\alpha = a_1 \wedge a_2 \wedge \cdots \wedge a_p, \tag{1}$$

$$\beta = b_1 \wedge b_2 \wedge \cdots \wedge b_q. \tag{2}$$

If A and B are the subspaces corresponding to α and β, respectively, and $B^\perp$ is the orthogonal complement of B in E^n, then we have

$$a_j = a_j^H + a_j^V,$$

where

$$a_j^H \in B, \quad a_j^V \in B^\perp, \quad j = 1, 2, \ldots, p.$$

Put

$$\begin{aligned} \alpha_H &= a_1^H \wedge a_2^H \wedge \cdots \wedge a_p^H, \\ \alpha_V &= a_1^V \wedge a_2^V \wedge \cdots \wedge a_p^V, \\ \alpha_M &= \alpha - \alpha_H - \alpha_V. \end{aligned}$$

Then we have a decomposition

$$\alpha = \alpha_H + \alpha_V + \alpha_M, \tag{3}$$

where α_H, α_V and α_M are called the horizontal, vertical and mixed parts, respectively, of α. By direct computation, we get the following lemmas.

LEMMA 2. *The decomposition* (3) *is invariant under transformations of basis in subspace A. Precisely, if in the linear subspace A we choose another basis* $\{a_i'\}$, *we have the corresponding p-vector* α' *and projections* α_H', α_V' *and* α_M', *and*

$$\alpha' = k\alpha, \quad \alpha_H' = k\alpha_H, \quad \alpha_V' = k\alpha_V, \quad \alpha_M' = k\alpha_M$$

for some nonzero scalar k.

LEMMA 3. $\langle \alpha, \alpha_H \rangle \geq 0$, *where equality holds if and only if* $\alpha_H = 0$.

LEMMA 4. $|\alpha|^2 = |\alpha_H|^2 + |\alpha_V|^2 + |\alpha_M|^2$.

Proof. One easily verifies that

$$\langle \alpha_H, \alpha_V \rangle = \langle \alpha_H, \alpha_M \rangle = \langle \alpha_V, \alpha_M \rangle = 0. \qquad \square$$

DEFINITION. The angle θ between subspaces A^p and $B^q, p \leq q$, is defined by

$$\theta = \arccos \frac{|\alpha_H|}{|\alpha|}. \tag{4}$$

Remark 1. The concept of the p-dimensional angle defined above is a natural generalization of classical angles such as the angles between two lines, a line and a plane, and between two planes. By Lemma 4, it is clear that θ is real and satisfies

$$0 \leq \theta \leq \pi/2;$$

further, $\theta = \pi/2$ if and only if the subspace A is orthogonal to B.

Remark 2. When $p = q$, formula (4) can be written as

$$\theta = \arccos \frac{|\langle \alpha, \beta \rangle|}{|\alpha| \cdot |\beta|}.$$

In this case, the p-dimensional angle θ between two decomposable p-vectors α and β over E^n is equal to the usual Euclidean angle between α and β as two vectors in the induced Euclidean space $\wedge^p E^n$.

3. First Properties

Using the projection α_H, we can easily generalize some well-known properties of usual angles to the higher dimensional case as follows.

THEOREM 1 (Reducibility). *Let A and B be subspaces of E^n with dimensions p and q, respectively, where $1 < p \leq q < n$. Suppose that $A \cap B \neq \{0\}$, and that the orthogonal complements of $A \cap B$ in A and in B are A' and B', respectively. Then the angle between A and B is equal to the angle between A' and B'.*

THEOREM 2 (Three cosines). *Suppose that $1 \leq p < q < n$, A^p and B^q are subspaces of E^n with dimensions p and q, respectively, and A^p is not orthogonal to B^q. Suppose C^p is the projection of A^p in B^q, and D^p is an arbitrary p-dimensional subspace of B^q. Denote by θ', θ and ϕ the angles between A^p and D^p, A^p and B^q, and C^p and D^p, respectively. Then*

$$\cos \theta' = \cos \theta \cos \phi.$$

Proof. Let α and δ be the p-vectors corresponding to A^p and D^p, respectively. Denote by α_H, α_V and α_M the horizontal, vertical and mixed parts, respectively, of α with respect to B^q. Then α_H corresponds to C^p, and we have

$$\langle \alpha_V, \delta \rangle = \langle \alpha_M, \delta \rangle = 0.$$

Hence

$$\begin{aligned}\cos\theta' &= \frac{|\langle\alpha,\delta\rangle|}{|\alpha|\cdot|\delta|} = \frac{|\langle\alpha_H,\delta\rangle|}{|\alpha|\cdot|\delta|} \\ &= \frac{|\alpha_H|}{|\alpha|}\cdot\frac{|\langle\alpha_H,\delta\rangle|}{|\alpha_H|\cdot|\delta|} \\ &= \cos\theta\cos\phi. \end{aligned}$$

□

COROLLARY (Minimum). *Suppose A^p and B^q are subspaces of E^n with dimensions p and q, respectively, $1 \leq p < q < n$. Then the angle θ between A^p and B^q equals the minimum of the angle between A^p and an arbitrary p-dimensional subspace D^p of B^q.*

4. Further Properties

Now we consider some more interesting properties of higher dimensional angles.

THEOREM 3 (Triangle inequality). *Let A, B and C be three different p-dimensional subspaces of $E^n, 1 < p < n$. Denote by θ_{AB}, θ_{AC} and θ_{BC} the angles between A and B, A and C, and B and C, respectively. Then*

$$\theta_{AB} + \theta_{BC} \geq \theta_{AC}. \tag{5}$$

Equality in (5) holds if and only if $\dim(A\cap B\cap C) = p-1$ and the 1-dimensional orthogonal complements A', B' and C' of $A\cap B\cap C$ in A, B and C, respectively, lie in one and the same plane and B' is placed in the smaller pair of vertical angles formed by A' and C'.

Proof. Denote $N = \dim \wedge^p E^n$. Let S^{N-1} be the unit hypersphere in $\wedge^p E^n$. Identifying opposite points of S^{N-1}, we obtain an elliptic space Σ^{N-1}. Three different p-dimensional subspaces A, B and C of E^n correspond to three different points A^*, B^* and C^* in Σ^{N-1}. Then θ_{AB}, θ_{BC} and θ_{AC} equal the distances between the points A^* and B^*, B^* and C^*, and A^* and C^*, respectively. Since Σ^{N-1} is a metric space (see, for example, Blumenthal [8]), inequality (5) is obtained.

For equality to hold in (5), it is necessary and sufficient that the point B^* lies on the metric segment A^*C^*. So there exist nonzero scalars h and k such that

$$\beta = h\alpha + k\gamma,$$

where α, β and γ are the decomposable p-vectors corresponding to the subspaces A, B and C, respectively. Since β is decomposable, we have

$$\beta \wedge y = 0, \quad y \in B,$$

and hence

$$h\alpha\wedge y + k\gamma\wedge y = 0, \quad y \in B. \tag{6}$$

Denoting

$$r = \dim(A \cap B \cap C),$$

from $\beta = h\alpha + k\gamma$ we get

$$\dim(A \cap C) = r.$$

Let A', B' and C' be the orthogonal complements of $A \cap B \cap C$ in A, B and C, respectively. Then

$$\dim A' = \dim B' = \dim C' = p - r,$$

$$A' \cap C' = A' \cap B' = B' \cap C' = \{0\}.$$

Denote by α', β' and γ' the $(p-r)$-vectors corresponding to A', B' and C', respectively. From (6) it follows that

$$h\alpha' \wedge y' + k\gamma' \wedge y' = 0, \quad y' \in B'.$$

One can choose a basis in E^n (which is not orthogonal in general) such that

$$\alpha' = a\, e_1 \wedge \cdots \wedge e_{p-r}, \quad \gamma' = c\, e_{p-r+1} \wedge \cdots \wedge e_{2p-2r},$$

where a and c are nonzero scalars. The vector y' can be decomposed as follows:

$$y' = \sum_{i=1}^{p-r} u_i e_i + \sum_{j=1}^{p-r} v_j e_{p-r+j} + \sum_{t=1}^{n-2p+2r} w_t e_{2p-2r+t}.$$

So we obtain

$$ha \sum_{j=1}^{p-r} v_j e_1 \wedge \cdots \wedge e_{p-r} \wedge e_{p-r+j} + kc \sum_{i=1}^{p-r} u_i e_{p-r+1} \wedge \cdots \wedge e_{2p-2r} \wedge e_i +$$

$$+ha \sum_{t=1}^{n-2p+2r} w_t e_1 \wedge \cdots \wedge e_{p-r} \wedge e_{2p-2r+t} +$$

$$+kc \sum_{t=1}^{n-2p+2r} w_t e_{p-r+1} \wedge \cdots \wedge e_{2p-2r} \wedge e_{2p-2r+t}$$

$$= 0.$$

If $r < p - 1$, then $p - r \geq 2$, we will get

$$hv_j = ku_i = hw_t = kw_t = 0$$

for all i, j and t. But $y' \neq 0$ and $A' \cap B' = B' \cap C' = \{0\}$, so we must have $h = k = 0$, which is impossible by assumption. Therefore we get

$$\dim(A \cap B \cap C) = p - 1.$$

Consequently, by Theorem 1, θ_{AB}, θ_{AC} and θ_{BC} are equal to the angles between A' and B', A' and C', and B' and C', where A', B' and C' are the 1-dimensional orthogonal complements of $A \cap B \cap C$ in A, B and C, respectively. The rest of the proof is trivial. □

THEOREM 4 (Complement). *Suppose A^p and B^q are subspaces of E^n with dimensions p and q, respectively, $1 \leq p < q < n, p \leq n - q$, $B^\perp$ the orthogonal complement of B in E^n. Denote by θ and $\theta^\perp$ the angles between A and B and between A and $B^\perp$, respectively. Then*

$$\cos^2 \theta + \cos^2 \theta^\perp \leq 1, \tag{7}$$

where equality holds if and only if either $p = 1$ or $\theta\theta^\perp = 0 (p > 1)$.

Proof. Inequality (7) follows from decomposition (3) and Lemma 4. Equality in (7) holds if and only if

$$\alpha_M = 0.$$

This occurs if and only if either $p = 1$, or $p > 1$ and $\alpha = \alpha_H$ or $\alpha = \alpha_V$. □

Theorems 3 and 4 show some of the differences between the cases of higher and lower dimensions.

5. Principal Angles

Let A^p and B^q be subspaces of E^n with dimensions p and q, respectively, $1 \leq p \leq q < n$. The principal angles between A^p and B^q, $0 \leq \theta_1 \leq \theta_2 \leq \cdots \leq \theta_p \leq \pi/2$, are given by

$$\begin{aligned} \cos\theta_i &= \frac{a_i \cdot b_i}{|a_i||b_i|} \\ &= \max\left\{\frac{a \cdot b}{|a||b|} : a \perp a_m, b \perp b_m, m = 1, 2, \ldots, i-1\right\}, \end{aligned}$$

where $a \in A^p, b \in B^q$ (see Miao and Ben-Israel [4, p. 81]).

The following theorem gives a simple relation between the higher dimensional angle θ and the principal angles $\theta_1, \theta_2, \ldots, \theta_p$.

THEOREM 5. *Suppose the angle between A^p and B^q is θ, and the principal angles between them are $\theta_1, \theta_2, \ldots, \theta_p$. Then*

$$\cos\theta = \cos\theta_1 \cos\theta_2 \cdots \cos\theta_p.$$

Proof. By normalization, all these a_i's and b_i's can be assumed to be unit vectors. Put $\alpha = a_1 \wedge a_2 \wedge \cdots \wedge a_p$, then α is the p-vector corresponding to subspace A^p, and $\langle \alpha, \alpha \rangle = 1$. Noting that

$$a_i^H = b_i \cos\theta_i, \quad i = 1, 2, \ldots, p,$$

we get

$$\alpha_H = \cos\theta_1 \cos\theta_2 \cdots \cos\theta_p b_1 \wedge b_2 \wedge \cdots \wedge b_p.$$

Therefore

$$\begin{aligned}\cos\theta &= |\alpha_H| \\ &= \cos\theta_1 \cos\theta_2 \cdots \cos\theta_p.\end{aligned}$$

□

Remark. The product of cosines of principal angles between subspaces L and M was denoted by $\cos\{L, M\}$ only as a symbol in [4]. Now, Theorem 5 shows that this symbol $\cos\{L, M\}$ is really the cosine of an angle.

6. Grassmann Manifolds

The set of all p-dimensional subspaces of E^n with suitable topology forms a Grassmann manifold $G(p, n-p)$. The theory of angles between subspaces of E^n is closely connected with the geometry of Grassmann manifolds. In this way, several theorems of distance geometry can be used to obtain the corresponding results of differential geometry.

For example, by normalization, we can restrict p-vectors to have unit lengths. All these unit p-vectors form a hypersphere S^{N-1} of $\wedge^p E^n$, where $N = \dim \wedge^p E^n = \binom{n}{p}$. Denote by G the submanifold of S^{N-1} consisting of points corresponding to decomposable unit p-vectors. Then G is the isometrically immersed image of the Grassmann manifold $G(p, n-p)$ into S^{N-1}. Research on geometric properties of G is an interesting topic in differential geometry (see, for example, W. H. Chen [9], who pointed out that G is an algebraic submanifold). In order to obtain the immersed equations for G into S^{N-1}, we choose an orthonormal basis $\{e_j\}$ in E^n, and then get an induced orthonormal basis $\{E_K\}$ in $\wedge^p E^n$, where

$$E_K = e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p}, \quad 1 \le i_1 < i_2 < \cdots < i_p \le n.$$

Therefore a unit p-vector α can be represented as

$$\alpha = \sum_K \lambda_K E_K, \quad \sum_K (\lambda_K)^2 = 1,$$

where

$$\lambda_K = u_{i_1 i_2 \cdots i_p}, \quad 1 \le i_1 < i_2 < \cdots < i_p \le n.$$

The quantities λ_K are the usual Cartesian coordinates in the induced Euclidean space E^N or $\wedge^p E^n$, and the $u_{i_1 i_2 \cdots i_p}$'s are called Grassmann coordinates of α; they will change sign if any two subscripts are interchanged. Our immersed equations are just the conditions for a p-vector to be simple, which are known as Plücker equations and can be written out explicitly as follows:

$$\sum_{j=1}^{p}(-1)^{j+1} u_{a_j b_2 \cdots b_p} u_{b_1 a_1 \cdots \hat{a}_j \cdots a_p} = u_{a_1 a_2 \cdots a_p} u_{b_1 b_2 \cdots b_p},$$

where subscripts $\{a_1, a_2, \ldots, a_p\}$ and $\{b_1, b_2, \ldots, b_p\}$ are two arbitrary arrangements of $\{1, 2, \ldots, p\}$, and $\hat{a}_j$ means that a_j is omitted. (See Jiang [10, p. 84]; for some particular cases also see Cartan [11, pp. 18–20]).

On the other hand, according to Menger, a subset of a metric space is called (metrically) convex provided it contains for each two of its points at least one between-point (see Blumenthal [8, p. 41]). Now from Theorem 3 and its proof we infer the following result.

THEOREM 6. *The Grassmann manifold $G(p, n-p)$ as a submanifold of a sphere is a metric space. When $n > 3$ and $p > 1$, it is not metrically convex.*

Theorem 6 shows another difference between higher and lower dimensions.

Acknowledgement

The author would like to thank the referee for many valuable suggestions and helpful criticisms.

References

1. Dörband, W.: Determinantensätze und Simplexeigenschaften, *Math. Nachr.* **44** (1970), 295–304.
2. Eriksson, F.: The law of sines for tetrahedra and n-simplices, *Geom. Dedicata* **7** (1978), 71–80.
3. Lin, S.-Y. T. and Lin, Y.-F.: The n-dimensional Pythagorean theorem, *Linear Multilinear Algebra* **26** (1990), 9–13.
4. Miao, J. and Ben-Israel, A.: On principal angles between subspaces in R^n, *Linear Algebra Appl.* **171** (1992), 81–98.
5. Hotelling, H.: Relations between two sets of variates, *Biometrica* **28** (1936), 321–377.
6. Bourbaki, N.: *Algèbre*, Hermann, Paris, 1948.
7. Flanders, H.: *Differential Forms*, Academic Press, New York, London, 1963.
8. Blumenthal, L. M.: *Theory and Applications of Distance Geometry*, Chelsea Pub. Co., New York, 1970.
9. Chen, W. H.: The differential geometry of Grassmann manifolds as submanifolds (in Chinese), *Acta Math. Sinica* **31** (1988), 46–53. (*MR* 89h: 53116.)
10. Jiang, S.: Identities of Laplace type and decomposable m-vectors, *Math. Practice Theory* (in Chinese), **1** (1993), 82–85. (*MR* 94d: 53026.)
11. Cartan, E.: *Les systèmes différentiels extérieurs et leurs applications géometriques*, Paris, 1945.