Distance and Parallelism Between Flats in $\mathbf{R}^n$

Arthur M. DuPré
Department of Mathematics
Rutgers University
Newark, New Jersey 07102

and

Seymour Kass
Engineering and Mathematics
University of Massachusetts at Boston
Boston, Massachusetts 02125

Submitted by Richard A. Brualdi

ABSTRACT

A simple matrix formula for the distance between two flats or affine spaces of $\mathbf{R}^n$ is derived. The geometry of intersection and of the degree of parallelism of two flats is elucidated.

PRELIMINARIES

Though the distance between two flats in $\mathbf{R}^n$ is a natural generalization of the distance between two points in $\mathbf{R}^n$, there does not seem to be a general distance formula in the literature. We will show that when flats are defined implicitly by matrix equations, a simple distance formula can be derived in terms of the defining matrices themselves. Additionally, the row and null spaces of the matrices help elucidate how parallelism and intersection work in $\mathbf{R}^n$.

Let a *flat* L, also known as an *affine subspace* of $\mathbf{R}^n$, be defined by the matrix equation $L = \{x \mid Dx = d\}$, where D has full row rank r_D. L is thus a coset of $N(D)$, the null space of its defining matrix D. $N(D)$ is the *direction space* of L, and by definition, $\dim L = \dim N(D)$.

© Elsevier Science Publishing Co., Inc., 1992
655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

Suppose we have two flats $L_1 = \{x | Ax = a\}$ and $L_2 = \{y | By = b\}$, where A is $r_A \times n$ and B is $r_B \times n$. Left multiplication by a matrix with zero null space leaves the defined space unchanged. In particular, multiplication by $(R^T)^{-1}$, from the QR factorization $A^T = QR$, produces an orthonormal-rowed defining matrix for L_1. Thus we may assume A and B have orthonormal rows, so that $AA^T = Ir_A$ and $BB^T = Ir_B$.

If L_1 and L_2 intersect, then $L_1 \cap L_2$ is a flat with direction space $N(A) \cap N(B)$ [4, pp. 16, 30], and $\dim(L_1 \cap L_2) = \dim L_1 + \dim L_2 - \dim[N(A) + N(B)]$.

The intersection of the row spaces of A and B, $\operatorname{Row} A \cap \operatorname{Row} B$, plays a role in this paper. In terms of null spaces,

$$\operatorname{Row} A \cap \operatorname{Row} B = N(A)^{\perp} \cap N(B)^{\perp} = [N(A) + N(B)]^{\perp}.$$

Let C be a matrix with $\operatorname{Row} C = \operatorname{Row} A \cap \operatorname{Row} B$. Then

$$\mathbf{R}^n = [N(A) + N(B)] \oplus \operatorname{Row} C. \tag{1}$$

If $C \neq 0$, we assume C has orthonormal rows.

THE DISTANCE FORMULA

Definition. The distance between L_1 and L_2 is $d(L_1, L_2) = \min\|x - y\|$, $x \in L_1$, $y \in L_2$.

Theorem 1.

$$d(L_1, L_2) = \|C(A^Ta - B^Tb)\| = \|C(x_0 - y_0)\|, \tag{2}$$

where x_0, y_0 are arbitrary elements of L_1 and L_2, respectively.

Proof. Case 1. $C = 0$. From (1), $N(A) + N(B) = \mathbf{R}^n$. That $L_1 \cap L_2 \neq \varnothing$ in this case is well known [4, p. 30]. We prove it for the sake of completeness.

If a and b are both zero, then $L_1 = N(A)$ and $L_2 = N(B)$ are vector spaces, and so meet. So let $a \neq 0$. Then $L_1 \neq N(A)$. Let $x \in L_1$, $x \notin N(A)$, and $y \in L_2$. Then $x + y \in \mathbf{R}^n = N(A) + N(B)$, so $x + y = x' + y'$ for some $x' \in N(A)$, $y' \in N(B)$. Then $x - x' = y' - y \neq 0$, since $x \neq x'$. But

$A(x-x')=a$, so $x-x' \in L_1$, and $B(y-y')=b$, so $y-y'=x'-x \in L_2$. Thus $x-x' \in L_1 \cap L_2 \neq \varnothing$. So the flats intersect.

Case 2. $C \neq 0$. C may be taken to have orthonormal rows by the following construction. Let

$$\begin{pmatrix} A \\ A^* \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} B \\ B^* \end{pmatrix}$$

be $n \times n$ orthogonal matrices, so that $AA^{*T}=0$, $BB^{*T}=0$. Since $N(A^*)=$ Row A, $N(B^*)=$ Row B, and Row $A \cap$ Row $B \neq 0$, any orthonormal basis of

$$N(A^*) \cap N(B^*) = N\begin{pmatrix} A^* \\ B^* \end{pmatrix}$$

will serve as the rows of C.

Every pair of points in L_1 determines a vector in $N(A)$: $N(A)=\{x-y \mid x, y \in L_1\}$. A vector is orthogonal to L_1 if it is orthogonal to all vectors determined by pairs of points in L_1. So all vectors orthogonal to L_1 lie in $N(A)^{\perp}=$ Row A. Similarly, all vectors orthogonal to L_2 lie in Row B. Thus all vectors orthogonal to L_1 and L_2 lie in Row $A \cap$ Row $B=$ Row C, and conversely.

A pair $p \in L_1$ and $q \in L_2$ is a *closest pair* if $\|q-p\|$ is minimal among all the distances between points of L_1 and L_2. This quantity is $d(L_1, L_2)$. We will show that, given $x \in L_1$ and $y \in L_2$, C^TC converts $x-y$ into a vector orthogonal to L_1 and L_2, representable as the difference of a closest pair of points.

Since C has orthonormal rows, $CC^T=I$, and (by [3, p. 158]) C^TC projects $\mathbf{R}^n$ onto Row C and $I-C^TC$ projects $\mathbf{R}^n$ onto the orthogonal complement $N(A)+N(B)$.

Let $x \in L_1$, $y \in L_2$. Then $C^TC(x-y)$ is representable as the difference of points in L_1 and L_2. For,

$$(I-C^TC)(x-y)=n_A+n_B, \qquad n_A \in N(A), \quad n_B \in N(B),$$

so

$$C^TC(x-y)=(x-y)-n_A-n_B=-(y+n_B)+(x-n_A),$$

where $y+n_B \in L_2$ and $x-n_A \in L_1$.

Moreover, $\|C(x-y)\|$ is minimal among all the distances between points of L_1 and L_2, because, with v_i^T the orthonormal rows of C,

$$\|C(y-x)\|^2 = \sum \left[v_i^T(y-x)\right]^2 \leqslant \|y-x\|^2,$$

by Bessel's inequality.

To complete the proof we show that $\|C(x-y)\|$ is the same for any choice of points $x \in L_1$, $y \in L_2$. Since $A^Ta \in L_1$ and $B^Tb \in L_2$, it suffices to show $\|C(x-y)\| = \|C(A^Ta - B^Tb)\|$.

A parametric representation of the flats is given by

$$\begin{aligned} Ax = a \quad &\Leftrightarrow \quad x = A^Ta + A^{*T}s, \quad s \in \mathbf{R}^{n-r_A}, \\ By = b \quad &\Leftrightarrow \quad y = B^Tb + B^{*T}t, \quad t \in \mathbf{R}^{n-r_B}. \end{aligned} \tag{3}$$

As s ranges over $\mathbf{R}^{n-r_A}$, $A^{*T}s$ ranges over the vectors in $N(A)$; similarly for $B^{*T}t$ and t.

Subtracting in (3),

$$x - y = A^Ta - B^Tb + A^{*T}s - B^{*T}t.$$

Applying C, using $CA^{*T} = CB^{*T} = 0$, we get $\|C(x-y)\| = \|C(A^Ta - B^Tb)\|$. Thus $d(L_1, L_2) = \|C(A^Ta - B^Tb)\|$, completing the proof. ■

PARALLELISM AND INTERSECTION

From Theorem 1 it is clear how $L_1 \cap L_2$ depends on C:

COROLLARY. *A necessary and sufficient condition for L_1 and L_2 to intersect is that $C = 0$ or $A^Ta - B^Tb \in N(C)$.*

If $L_1 \cap L_2 = 0$ and $\dim[N(A) \cap N(B)] = k$, then L_1, L_2 are said to have *degree of parallelism* k. If $k = 0$, the flats are *totally skew*, i.e., have no common direction. At $k = 1$, the flats have a line of common direction, etc. Full parallelism occurs when the direction space of one is contained in the direction space of the other.

The subspaces of common direction are the *end spaces*, E_i.

DEFINITION.

$$E_1 = \{x \in L_1 | d(x, L_2) = d(L_1, L_2)\},$$

$$E_2 = \{y \in L_2 | d(L_1, y) = d(L_1, L_2)\}.$$

For nonintersecting flats we seek the defining equations of E_1, E_2; for intersecting flats, the equation of $L_1 \cap L_2$.

LEMMA 1. *If* $L \subseteq L'$ *and* $\dim L = \dim L'$, *then* $L = L'$.

Proof [1, *Chapter* 3]. An exercise in cosets. Say $L = u + N(D)$ and $L' = u' + N(D')$. First, show $L \subseteq L'$ implies $N(D) \subseteq N(D')$. Then $\dim L = \dim L'$ implies $\dim N(D) = \dim N(D')$, and the usual vector-subspace argument gives $N(D) = N(D')$. Thus $L = L'$. ∎

LEMMA 2. *If* $L = \{x | Dx = d\}$ *and* $L' = \{x | D^T Dx = D^T d\}$, *then* $L = L'$.

Proof. Clearly, $L \subseteq L'$. Since $N(D) = N(D^T D)$ [5, p. 157], $\dim L = \dim L'$. Then $L = L'$ by Lemma 1. ∎

LEMMA 3. $A^T A C^T C = C^T C$ *and* $B^T B C^T C = C^T C$.

Proof. Identically, for all $x \in \mathbf{R}^n$,

$$(A^T A C^T C - C^T C)x = A^T A(C^T Cx) - (C^T Cx) = (A^T A - I)(C^T Cx).$$

Now $A^T A(C^T Cx) \in \operatorname{Row} A$ and $(C^T Cx) \in \operatorname{Row} A \cap \operatorname{Row} B$, so $(A^T A C^T C - C^T C)x \in \operatorname{Row} A$. However, $A^T A - I$ maps $\mathbf{R}^n$ onto $N(A)$, so $(A^T A - I)(C^T Cx) \in N(A)$. Since $N(A) \cap \operatorname{Row} A = 0$, we have $(A^T A C^T C - C^T C)x = 0$, or $A^T A C^T C = C^T C$. Similarly, $B^T B C^T C = C^T C$. ∎

REMARK. From Lemma 3,

$$C^T C = \tfrac{1}{2}(A^T A + B^T B) C^T C. \tag{4}$$

THEOREM 2. *The equations of the end spaces E_1 and E_2 are, respectively,*

$$(A^TA + B^TB)x = A^Ta + B^Tb + C^TC(A^Ta - B^Tb), \tag{5}$$

$$(A^TA + B^TB)x = A^Ta + B^Tb - C^TC(A^Ta - B^Tb). \tag{5'}$$

If L_1 and L_2 intersect, the equation of $L_1 \cap L_2$ is

$$(A^TA + B^TB)x = A^Ta + B^Tb. \tag{6}$$

Proof. We prove (5). The proof of (5′) is similar. For any flat $L = \{z | Dz = d\}$, where $DD^T = I$, the orthogonal projection of any $x \in \mathbf{R}^n$ onto L is $z = (I - D^TD)x + D^Td$ [2, p. 94]. Thus $A^TA(z - x) = A^Td - A^TAx$ is a vector from x to $z \in L$, orthogonal to L. If $x \in L$, then $z = x$ and $A^Td - A^TAx = 0$.

Now let $x \in L_1$ and $L = L_2$. The orthogonal projection of $x \in L_1$ onto L_2 is the point $z = (I - B^TB)x + B^Tb \in L_2$, and $x - z = B^TBx - B^Tb$ is the vector from x to z, which is orthogonal to L_2. If x and z are in the same space, so that L_1 and L_2 intersect, then $C^TC(A^Ta - B^Tb) = 0$ in what follows.

In order for x and z to be a closest pair, we must have

$$x - z = B^TBx - B^Tb = C^TC(A^Ta - B^Tb) \quad \textit{and} \quad Ax = a.$$

Thus $E_1 = \{x | Ax = a\} \cap \{x | B^TBx = B^Tb + C^TC(A^Ta - B^Tb)\}$.

Since, by Lemmas 1 and 2,

$$\{x | B^TBx = B^Tb + C^TC(A^Ta - B^Tb)\} = \{x | Bx = b + BC^TC(A^Ta - B^Tb)\},$$

we have

$$E_1 = \{x | Ax = a\} \cap \{x | Bx = b + BC^TC(A^Ta - B^Tb)\}.$$

Thus E_1 is defined by

$$\begin{pmatrix} A \\ B \end{pmatrix} x = \begin{pmatrix} a \\ b + BC^TC(A^Ta - B^Tb) \end{pmatrix}.$$

Apply $\begin{pmatrix} A \\ B \end{pmatrix}^T$ to both sides to get

$$(A^TA + B^TB)x = A^Ta + B^Tb + B^TBC^TC(A^Ta - B^Tb).$$

By Lemma 3, $B^TBC^TC(A^Ta - B^Tb) = C^TC(A^Ta - B^Tb)$. Thus the equation of E_1 is $(A^TA + B^TB)x = A^Ta + B^Tb + C^TC(A^Ta - B^Tb)$. Similarly for E_2.

The spaces intersect if and only if $C^TC(A^Ta - B^Tb) = 0$, from which (6) follows. ∎

EXAMPLE 1 (Distance between two points). In this case both A and B are invertible, so $A^* = 0$ and $B^* = 0$, and $N(A^*) = N(B^*) = \mathbf{R}^n$. Then, from (3), $N(A^*) \cap N(B^*) = \mathbf{R}^n$, and so we may choose $C = I$, which has row space $\mathbf{R}^n$. Since $A^T = A^{-1}$ and $B^T = B^{-1}$,

$$d(L_1, L_2) = \|C(A^Ta - B^Tb)\| = \|I(A^{-1}a - B^{-1}b)\| = \|x - y\|.$$

EXAMPLE 2. The distance from point y to a flat $L: Ax = a$ is

$$d(y, L) = \|Ay - a\|. \tag{7}$$

Proof: In (2), B is invertible and L_2 consists of the single point y. Then $N(A^*) \cap N(B^*) = N(A^*) \cap \mathbf{R}^n = N(A^*)$, so we may take $C = A$. Then $d(L_1, L_2) = d(L_1, y) = \|C(A^Ta - B^Tb)\| = \|A(A^Ta - B^{-1}b)\| = \|A(A^Ta - y)\| = \|a - Ay\| = \|Ay - a\|$.

Formula (7) gives the distance from a point to a hyperplane in $\mathbf{R}^n$. For an alternate proof of (7) see [2, p. 94]

EXAMPLE 3 (Lines in $\mathbf{R}^3$). Let lines L_1 and L_2 be defined by $Ax = a$ and $Bx = b$, where each 2×3 matrix

$$A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

has orthonormal rows. Then $\dim N(A) = \dim N(B) = 1$, so that either $N(A) \cap N(B) = 0$ or $N(A) = N(B)$.

Intersecting lines. By the corollary, flats meet if and only if $C = 0$ or $A^Ta - B^Tb \in N(C)$. Since we are in 3-space, $C \neq 0$; thus the lines meet if and only if $A^Ta - B^Tb \in N(C)$.

If $N(A) \cap N(B) = 0$, the lines meet in a point, since $\dim(N(A) \cap N(B)) = 0$. In this case $N(A^TA + B^TB) = N(A) \cap N(B)$, so $A^TA + B^TB$ is invertible, and from (6), the intersection point is $(A^TA + B^TB)^{-1}(A^Ta + B^Tb)$. If $N(A) \equiv N(B)$, the lines are identical, since $\dim N(A) = \dim N(B) = 1$.

Nonintersecting lines. Then $A^Ta - B^Tb \notin N(C)$. Again, either $N(A) \cap N(B) = 0$ or $N(A) = N(B)$. If $N(A) \cap N(B) = 0$, the end spaces consist of single points and the lines are skew. We use the general formula (2) to derive the standard formula for the distance between skew lines in $\mathbf{R}^3$. Since $\dim N(A) = \dim N(B) = 1$, we have $N(A) = \langle a_1 \times a_2 \rangle \neq 0$, $N(B) = \langle b_1 \times b_2 \rangle \neq 0$, and $\{a_1 \times a_2, b_1 \times b_2\}$ is linearly independent. Thus C is the matrix with the single orthonormal row $(a^* \times b^*)/\|a^* \times b^*\|$, where $a^* = a_1 \times a_2$ and $b^* = b_1 \times b_2$. Substituting into (2), we get

$$d(L_1, L_2) = \frac{\left\|(a^* \times b^*)^T(A^Ta - B^Tb)\right\|}{\|a^* \times b^*\|} = \frac{\left\|(a^* \times b^*)^T(x_0 - y_0)\right\|}{\|a^* \times b^*\|}$$

for any $x_0 \in L_1$, $y_0 \in L_2$. This is the standard distance formula [6, p. 336]. The end-space points are $(A^TA + B^TB)^{-1}[A^Ta + B^Tb \pm C^TC(A^Ta - B^Tb)]$. Using (4), it follows at once that the distance between them is $\|C(A^Ta - B^Tb)\|$.

Parallel lines. If $N(A) = N(B)$, then $\{a_1 \times a_2, b_1 \times b_2\}$ is linearly dependent and so $C = A$ or B. If $C = A$, then $d(L_1, L_2) = \|A(A^Ta - B^Tb)\| = \|a - A(B^Tb)\| = \|A(B^Tb) - a\|$, the distance from point $B^Tb \in L_2$ to L_1. If $C = B$, then $d(L_1, L_2) = \|B(A^Ta) - b\|$, the distance from point $A^Ta \in L_1$ to L_2.

It is now clear how intersection and parallelism of lines in $\mathbf{R}^3$ depend on A, B, and C:

Intersection	$L_1 = L_2$	Skew	Parallel
$A^Ta - B^Tb \in N(C)$	$A^Ta - B^Tb \in N(C)$	$A^Ta - B^Tb \notin N(C)$	$A^Ta - B^Tb \notin N(C)$
$N(A) \cap N(B) = 0$	$N(A) = N(B)$	$N(A) \cap N(B) = 0$	$N(A) = N(B)$

EXAMPLE 4 (Planes in $\mathbf{R}^4$). Let planes L_1 and L_2 be defined by $Ax = a$ and $Bx = b$, where A and B are 2×4 matrices with orthonormal rows. There are five possibilities.

Intersecting planes. $C = 0$ or $A^Ta - B^Tb \in N(C)$:

(a) $C = 0 \Rightarrow \text{Row}\, A \cap \text{Row}\, B = 0 \Rightarrow N(A) + N(B) = \mathbf{R}^4 \Rightarrow \dim[N(A) \cap N(B)] = 0 \Rightarrow L_1 \cap L_2$ is a point.

(b) dim Row $C = 1$ and $A^T a - B^T b \in N(C) \Rightarrow \dim[N(A) + N(B)] = 3 \Rightarrow \dim[N(A) \cap N(B)] = 1 \Rightarrow L_1 \cap L_2$ is a line.
(c) dim Row $C = 2$ and $A^T a - B^T b \in N(C) \Rightarrow N(A) = N(B) \Rightarrow L_1 = L_2$.

Nonintersecting planes. $C \neq 0$ and $A^T a - B^T b \notin N(C)$:

(d) dim Row $C = 1$ and $A^T a - B^T b \notin N(C) \Rightarrow \dim[N(A) + N(B)] = 3 \Rightarrow \dim[N(A) \cap N(B)] = 1 \Rightarrow E_1$ and E_2 are parallel lines, i.e., L_1 and L_2 are parallel of degree 1.
(e) dim Row $C = 2$ and $A^T a - B^T b \notin N(C) \Rightarrow N(A) = N(B) \Rightarrow L_1$ and L_2 are fully parallel.

There are no skew planes in spaces of dimension less than 5. In general, from (1) and $C \neq 0$, it is easy to show that a necessary condition for nonintersecting flats to be skew in $\mathbf{R}^n$ is that $\dim L_1 + \dim L_2 < n$. Similarly, one can show that full parallelism occurs in spaces of all dimensions.

REMARKS.

(1) The proof of Theorem 1 draws on the classical proof in [3].

(2) If A and B have full row rank, but not orthonormal rows, then our results remain valid with the Moore-Penrose inverses, A^+ and B^+, in place of A^T and B^T.

REFERENCES

1 David M. Bloom, *Linear Algebra and Geometry*, Cambridge U. P., 1979.
2 Seymour Kass, Spaces of closest fit, *Linear Algebra Appl.* 117:93–97 (1989).
3 O. Schreier and E. Sperner, *Modern Algebra and Matrix Theory*, Chelsea, New York, 1959.
4 Earnest Snapper and Robert J. Troyer, *Metric Affine Geometry*, Academic, 1971.
5 Gilbert Strang, *Linear Algebra and Its Applications*, Harcourt Brace Jovanovich, San Diego, 1988.
6 W. Wooten, E. F. Beckenbach, and F. J. Fleming, *Modern Analytic Geometry*, Houghton Mifflin, Boston, 1972.

Received 20 September 1989; final manuscript accepted 5 June 1991