7.1 Orthogonal Matrices

DEFINITION 1 A square matrix A is said to be ***orthogonal*** if its transpose is the same as its inverse, that is, if

$$A^{-1} = A^T$$

or, equivalently, if

$$AA^T = A^TA = I \tag{1}$$

THEOREM 7.1.1 *The following are equivalent for an $n \times n$ matrix A.*

(*a*) *A is orthogonal.*

(*b*) *The row vectors of A form an orthonormal set in R^n with the Euclidean inner product.*

(*c*) *The column vectors of A form an orthonormal set in R^n with the Euclidean inner product.*

THEOREM 7.1.2

(*a*) *The transpose of an orthogonal matrix is orthogonal.*

(*b*) *The inverse of an orthogonal matrix is orthogonal.*

(*c*) *A product of orthogonal matrices is orthogonal.*

(*d*) *If A is orthogonal, then $\det(A) = 1$ or $\det(A) = -1$.*

THEOREM 7.1.3 *If A is an $n \times n$ matrix, then the following are equivalent.*

(*a*) *A is orthogonal.*

(*b*) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ *for all* $\mathbf{x}$ *in* R^n.

(*c*) $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ *for all* $\mathbf{x}$ *and* $\mathbf{y}$ *in* R^n.

If A is an orthogonal matrix and T_A: $R^n \rightarrow R^n$ is multiplication by A, then we will call T_A an ***orthogonal operator*** on R^n. It follows from parts (*a*) and (*b*) of Theorem 7.1.3 that the orthogonal operators on R^n are precisely those operators that leave the lengths (norms) of vectors unchanged. However, this implies that orthogonal operators also leave angles and distances between vectors in R^n unchanged since these can be expressed in terms of norms

THEOREM 7.1.4 *If S is an orthonormal basis for an n-dimensional inner product space V, and if*

$$(\mathbf{u})_S = (u_1, u_2, \ldots, u_n) \quad and \quad (\mathbf{v})_S = (v_1, v_2, \ldots, v_n)$$

then:

(a) $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$

(b) $d(\mathbf{u}, \mathbf{v}) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$

(c) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$

THEOREM 7.1.5 *Let V be a finite-dimensional inner product space. If P is the transition matrix from one orthonormal basis for V to another orthonormal basis for V, then P is an orthogonal matrix.*

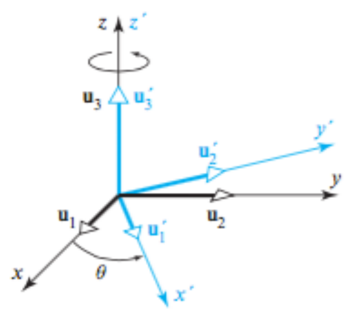
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

or, equivalently,

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned}$$

These are sometimes called the ***rotation equations*** for R^2.

Application to Rotation o fAxes in 3-Space



$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

A linear operator on R^2 is called ***rigid*** if it does not change the lengths of vectors, and it is called ***angle preserving*** if it does not change the angle between nonzero vectors.

7.2 Orthogonal Diagonalization

In this section we will be concerned with the problem of diagonalizing a symmetric matrix A. As we will see, this problem is closely related to that of finding an orthonormal basis for R^n that consists of eigenvectors of A. Problems of this type are important because many of the matrices that arise in applications are symmetric.

DEFINITION 1 If A and B are square matrices, then we say that B is ***orthogonally similar*** to A if there is an orthogonal matrix P such that $B = P^TAP$.

If A is orthogonally similar to some diagonal matrix, say $P^TAP = D$, then we say that A is orthogonally diagonalizable and that P orthogonally diagonalizes A

$A^T = (PDP^T)^T = (P^T)^TD^TP^T = PDP^T = A$ so A must be symmetric if it is orthogonally diagonalizable.

THEOREM 7.2.1 *If A is an $n \times n$ matrix with real entries, then the following are equivalent.*

(*a*) *A is orthogonally diagonalizable.*

(*b*) *A has an orthonormal set of n eigenvectors.*

(*c*) *A is symmetric.*

THEOREM 7.2.2 *If A is a symmetric matrix with real entries, then:*

(*a*) *The eigenvalues of A are all real numbers.*

(*b*) *Eigenvectors from different eigenspaces are orthogonal.*

Orthogonally Diagonalizing an $n \times n$ Symmetric Matrix

Step 1. Find a basis for each eigenspace of A.

Step 2. Apply the Gram–Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

Step 3. Form the matrix P whose columns are the vectors constructed in Step 2. This matrix will orthogonally diagonalize A, and the eigenvalues on the diagonal of $D = P^TAP$ will be in the same order as their corresponding eigenvectors in P.

Spectral Decomposition

If A is a symmetric matrix that is orthogonally diagonalized by $P = [\mathbf{u}_1\ \mathbf{u}_2 \cdots \mathbf{u}_n]$ and if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A corresponding to the unit eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$, then we know that $D = P^T AP$, where D is a diagonal matrix with the eigenvalues in the diagonal positions. $\mathbf{u}\mathbf{u}^T$ is the standard matrix for the orthogonal projection of R^n on the subspace spanned by the vector **u**.

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

The spectral decomposition of A tells that the image of a vector **x** under multiplication by a symmetric matrix A can be obtained by projecting **x** orthogonally on the lines (one-dimensional subspaces) determined by the eigenvectors of A, then scaling those projections by the eigenvalues, and then adding the scaled projections.

Schur decomposition of A

THEOREM 7.2.3 Schur's Theorem

If A is an $n \times n$ matrix with real entries and real eigenvalues, then there is an orthogonal matrix P such that P^TAP is an upper triangular matrix of the form

$$P^TAP = \begin{bmatrix} \lambda_1 & \times & \times & \cdots & \times \\ 0 & \lambda_2 & \times & \cdots & \times \\ 0 & 0 & \lambda_3 & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \quad (11)$$

in which $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A repeated according to multiplicity.

That is every square matrix A is orthogonally similar to an upper triangular matrix that has the eigenvalues of A on the main diagonal.

$A = PSP^T$ which is called a **Schur decomposition of A**

Upper Hessenberg decomposition

THEOREM 7.2.4 Hessenberg's Theorem

If A is an $n \times n$ matrix with real entries, then there is an orthogonal matrix P such that P^TAP is a matrix of the form

$$P^TAP = \begin{bmatrix} \times & \times & \cdots & \times & \times & \times \\ \times & \times & \cdots & \times & \times & \times \\ 0 & \times & \ddots & \times & \times & \times \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \times & \times & \times \\ 0 & 0 & \cdots & 0 & \times & \times \end{bmatrix} \tag{13}$$

Every square matrix with real entries is orthogonally similar to a matrix in which each entry below the first ***subdiagonal*** is zero

$A = PHP^T$ which is called an ***upper Hessenberg decomposition*** of A.