Orthogonal matrices

Definition: A square matrix A is called *orthogonal* if it is invertible and $A^{-1} = A^T$.

Note: Clearly A is orthogonal iff and only if $A\,A^T = I$ or if and only if $A^T A = I$.

Example: Rotation by an angle θ matrices $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ are orthogonal matrices.

Reflection matrices (about coordinate axes or other lines) are also orthogonal: $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Why do we call such matrices orthogonal? Because their columns (and rows) are, in fact, orthogonal vectors (actually even orthonormal).

Theorem: A square matrix is orthogonal if and only if its row vectors form an orthonormal set (with respect to the Euclidian inner product). The same holds for its column vectors.

Orthogonal matrices

<u>Theorem</u> (basic properties of orthogonal matrices):

a) If A is an orthogonal matrix, then A^{-1} is also orthogonal.

b) If A and B are orthogonal matrices, then $A \cdot B$ is also orthogonal.

c) If A is an orthogonal matrix, then $\det(A) = 1$ or $\det(A) = -1$.

<u>Theorem</u> (preservation of length and dot product): For a square matrix A, the following statements are equivalent:

a) A is orthogonal.

b) $\|Ax\| = \|x\|$ for all vectors x in $\mathbb{R}^n$.

c) $Ax \cdot Ay = x \cdot y$ for all vectors x, y in $\mathbb{R}^n$.

<u>Reminder from MA1201</u>: Let u, v be any vectors in $\mathbb{R}^n$ and let A be any $n \times n$ matrix. The following two properties of the dot product in $\mathbb{R}^n$ are crucial in the proof of the above theorem.

- $Au \cdot v = u \cdot A^T v$
- $u \cdot v = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$

Change of orthonormal basis

<u>Theorem</u>: Let V be a finite dimensional inner product space and let B be an orthonormal basis in V.
If u and v are vectors in V and $(u)_B = (u_1, u_2, \ldots, u_n)$, $(v)_B = (v_1, v_2, \ldots, v_n)$ are their representations in the basis B, then

a) $\langle u, v \rangle = u_1\, v_1 + u_2\, v_2 + \ldots + u_n\, v_n$

b) $\|u\| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$
In other words, $\|u\|_V = \|(u)_B\|_{\mathbb{R}^n}$.

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

<u>Theorem</u>: The transition matrix $P_{B' \to B}$ from an orthonormal basis B' to another orthonormal basis B is orthogonal.

Change of orthonormal basis

Example: Given a rectangular coordinate system $x - y$, consider the coordinate system $x' - y'$ obtained by rotating the x and y axes by an angle θ.

Then the coordinates (x', y') of a point Q in the new coordinate system are obtained from the old coordinates (x, y) by multiplication with the rotation matrix R_θ^T. In other words,

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

That is because if $B = \{u_1, u_2\}$ is the orthonormal basis obtained by choosing a unit vector u_1 on the x axis and a unit vector u_2 on the y-axis, and if $B' = \{u_1', u_2'\}$ is the orthonormal basis obtained by rotating u_1 and u_2 by the angle θ, then the transition matrices between these two bases are

$$P_{B' \to B} = R_\theta$$

and

$$P_{B \to B'} = R_\theta^{-1} = R_\theta^T$$

Orthogonal diagonalization

Definition: A matrix A is called *diagonalizable* if there exists an invertible matrix P and a diagonal matrix D such that

$$P^{-1} A P = D$$

A matrix A is called *orthogonally diagonalizable* if the invertible matrix P that diagonalizes it is in fact *orthogonal*. In other words,

$$P^T A P = D$$

Theorem: For an $n \times n$ matrix A, the following are equivalent:

a) A is *orthogonally* diagonalizable.
b) A has an *orthonormal* set of n eigenvectors.
   In other words, there is an orthonormal basis in $\mathbb{R}^n$ consisting of eigenvectors of A.
c) A is *symmetric*.

Theorem: If A is a symmetric matrix, then

- Its eigenvalues are all real numbers.
- Eigenvectors corresponding to different eigenvalues are orthogonal.

Procedure to orthogonally diagonalize a matrix

Let A be a square matrix.

Step 0: Verify that A is symmetric.
If it is symmetric, then it is orthogonally diagonalizable.
If it is not symmetric, stop (since then it cannot be orthogonally diagonalized).

Step 1: Find the eigenvalues of A, and for each eigenvalue, find a basis for the corresponding eigenspace.

Step 2: Apply Gram-Schmidt to each of these bases to obtain an ONB for each eigenspace.
Note that putting together all of these bases for each eigenspace leads to an ONB for the whole space.

Step 3: Form the matrix P whose columns are the vectors obtained in Step 2 (mind the order). This matrix is orthogonal and it diagonalizes A.
Moreover, let D be the diagonal matrix whose diagonal entries are the eigenvalues of A (mind the order). Then

$$P^T A P = D$$

The adjoint matrix

Definition: The *adjoint* (or conjugate transpose) of a matrix A is the matrix A^* defined as

$$A^* = \overline{A^T} = \overline{A}^T$$

Theorem (basic properties of the transpose matrix):

a) $(A^*)^* = A$

b) $(A + B)^* = A^* + B^*$

c) $(k\,A)^* = \overline{k}\,A^*$

d) $(A\,B)^* = B^*\,A^*$

Note: Given two vectors u, v in $\mathbb{C}^n$, since $u \cdot v = \overline{v}^T\,u$, we can now write

$$u \cdot v = v^*\,u$$

Hermitian and unitary matrices

Definition: A square matrix A is called *self-adjoint* (or Hermitian) if

$$A^* = A$$

A square matrix A is called *unitary* if it is invertible and

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

Note: The concept of self-adjoint (Hermitian) matrix is the analogue for complex matrices of the concept of symmetric matrix. Similarly, the concept of unitary matrix is the analogue for complex matrices of the concept of orthogonal matrix.

Theorem: If A is a Hermitian matrix, then

- Its eigenvalues are all real numbers.
- Eigenvectors corresponding to different eigenvalues are orthogonal.

Unitary matrices, unitary diagonalization

Theorem: Given an $n \times n$ complex matrix, the following statements are equivalent:

a) A is unitary.
b) $\|Ax\| = \|x\|$ for all x in $\mathbb{C}^n$.
c) $Ax \cdot Ay = x \cdot y$ for all x, y in $\mathbb{C}^n$.
d) The column vectors of A form an orthonormal basis in $\mathbb{C}^n$. Similarly for the row vectors.

Definition: A square matrix A is *unitarily diagonalizable* if there are a unitary matrix P and a diagonal matrix D such that

$$P^* A P = D$$

Theorem: If A is Hermitian, then A is unitarily diagonalizable and $P = [c_1, c_2, \ldots, c_n]$, where $c_1, c_2, \ldots, c_n$ form an orthonormal set of eigenvectors for A, while D has the eigenvalues of A as diagonal entries.

The procedure for unitary diagonalization is the same as for orthogonal diagonalization.

Normal matrices

Question: Is everything exactly the same for complex matrices as for real matrices?

NO. If A is unitarily diagonalizable, then A is not always Hermitian. However, it has to satisfy a more general (although related property: $A\,A^* = A^*A$.

Definition: A square matrix A is called *normal* if

$$A\,A^* = A^*A$$

Theorem: A square matrix is normal iff it is unitarily diagonalizable.