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

Theorem (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 $\det(A) = -1$.

Theorem (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 $\mathbb{R}^n$.

Reminder from MA1201: 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

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

Theorem: 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

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