Caleb Powell

Dr. Gates

MAT-223 Linear Algebra

December 2 2019

Orthogonal Matrices

An orthogonal matrix is an interesting case of an $n \times n$ matrix A that satisfies the equation $A^T A = I$. These types of matrices have very many interesting characteristics to them. When viewed as linear transformations such as reflections, rotations, and compositions of reflections and rotations on $\mathbb{R}^n$, they preserve lengths and angles. Their determinants and eigenvalues are always either positive or negative 1, their column vectors form an orthonormal set in $\mathbb{R}^n$, and they can be viewed as a group in Abstract Algebra. Thus, these matrices are really nice in terms of geometry.

To provelength conservation, let's first suppose that a vector $\vec{x}$ undergoes a rotational transformation L. If the length is preserved, then the magnitude of the vector should be the same before and after the linear transformation such that $\|\vec{x}\|^2 = \|L\vec{x}\|^2$. Since the vector $\vec{x}$ is a $n \times 1$ matrix and the square of the magnitude of the vector $\vec{x}$ is equivalent to the dot product ($\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$), we know that the dot product is equal to $x^T x$ since,

$x \cdot x = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x^T x$. With this information, we can find the square magnitude of the rotational transformation of $\vec{x}$ to be $\|L\vec{x}\|^2 = L\vec{x} \cdot L\vec{x} = (L\vec{x})^T L\vec{x}$. Since we also know that you can distribute the transpose of the product of two matrices such that $(AB)^T = B^T A^T$, we know that: $(L\vec{x})^T L\vec{x} = x^T L^T L x$. Given that the rotational transformation L is an Orthogonal Matrix, we know that $L^T L = I$ and thus, $L^T = L^{-1}$. From this we have all the information we need to show that: $\|L\vec{x}\| = x^T L^T L x = x^T (L^{-1} L) x = x^T x = x \cdot x =$ $\boxed{\|x\| = \|L\vec{x}\|}$. This proves that when viewed as a linear transformation, Orthogonal Matrices preserve length.

To prove that angles are also preserved we must first remember that the dot product of two vectors $\vec{x}$ and $\vec{y}$ is defined as $\vec{x} \cdot \vec{y} = \|\vec{x}\|\|\vec{y}\| cos\theta$. From here, we can find the equation for calculating the angle between two vectors to be $cos\theta = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|}$. When we apply an orthogonal transformation L to $\vec{x}$ and $\vec{y}$ we find that the angels are in fact preserved since:

$$cos\theta_L = \frac{L\vec{x} \cdot L\vec{y}}{\|L\vec{x}\|\|L\vec{y}\|} = \frac{(L\vec{x})^T L\vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \frac{\vec{x}^T L^T L\vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \frac{\vec{x}^T \vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|\|\vec{y}\|} = cos\theta$$

.

Orthogonal matrices also have a lot of nice properties in terms of what determinants they can have. The determinants of an Orthogonal matrix must all be either positive or negative one. We can see that this is true if we take the determinant of both sides of the equation $Q^T Q = I$. Because we know that the determinant of the identity matrix is equal to 1, then it must be true that $det(Q^T Q) = det(I) = 1$. We also know that we can break the determinant of the product of two matrices into the the product of the determinant of each matrix such that $det(Q^T)det(Q) = 1$. Since it is true that $det(A) = det(A^T)$, we can rewrite this equation to be $det(Q)^2 = 1$. If we sqaure root both sides of this equation we find that $det(Q) = \pm 1$, where Q is an orthogonal matrix.

To determine what Eigenvalues Orthogonal matrices can have, we must recall that that an eigenvalue is some scalar λ such that $Ax = \lambda x$. We can then multiply both sides of this equation by $(Ax)^T$ to find that $(Ax)^T Ax = (Ax)^T \lambda x$. Given that $Ax = \lambda x$, this equation can then be rewritten so that $(Ax)^T Ax = (\lambda x)^T \lambda x$. We can then distribute the transpose to find that $x^T A^T Ax = x^T \lambda^T \lambda x$. Because A is orthogonal, $A^T A = I$, and since λ is a scalar, its also true that $\lambda^T \lambda = \lambda^2$. Therefore, we can write our equation to be that $x^T x = x^T \lambda^2 x$. Because $x^T x$ is a scalar value, it can be divide from the right side of the equation so that we have $\frac{x^T x}{x^T x} = \lambda^2$. We now can simplify to find that $\lambda^2 = 1$. If we square root both sides, we will find that $\lambda = \pm 1$ for all Orthogonal Matrices.

From the definition of what an orthogonal matrix is, we know that an $n \times n$ matrix A is said to be Orthogonal if the column vectors of A forms an orthonormal set in $\mathbb{R}^n$. An orthonormal set is defined such that $\vec{x_i} \cdot \vec{x_j} = \vec{x_i}^T \vec{x_j} = \begin{Bmatrix} 1 & \text{if } i = j \\ 0 & \text{if } j \neq i \end{Bmatrix}$. From the definition of the dot product being $\vec{x_i} \cdot \vec{x_j} = \|\vec{x_i}\| \|\vec{x_j}\| cos\theta$, we know that $\|\vec{x_i}\| = 1$ for all column vectors $i = 1, 2, \cdots, k$ of the matrix A. This also means that one column vector of A must be orthogonal (perpendicular) to some other column vector of A given that $cos(90) = 0$ and thatn when $i \neq j$, $\vec{x_i}^T \vec{x_j} = 0$.

Take for example the orthogonal matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. We can see that this is an orthogonal matrix since $A^T A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = I$. Since we know this is an Orthogonal matrix, if we take the magnitude of each of the column vectors $\vec{a_i}$, they should be equal to one.

$$\|\vec{a_1}\|^2 = \vec{a_1} \cdot \vec{a_1} = \vec{a_1}^T \vec{a_1} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = (0^2 + 0^2 + 1^2) = 1$$

$$\|\vec{a_2}\|^2 = \vec{a_2} \cdot \vec{a_2} = \vec{a_2}^T \vec{a_2} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = (1^2 + 0^2 + 0^2) = 1$$

$$\|\vec{a_3}\|^2 = \vec{a_3} \cdot \vec{a_3} = \vec{a_3}^T \vec{a_3} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = (0^2 + 1^2 + 0^2) = 1$$

This is exactly what we would have expected to get! We would also expect that each column vector of A would be orthogonal to some other column vector of A. We can see this is true with all of the column vectors of the example matrix A. For example:

$$\vec{a_1}^T\vec{a_3} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = (0^2 + 0(1) + 1(0) = 0$$

implies that the column vector $\vec{a_1}$ is orthogonal (perpendicular) to the column vector $\vec{a_3}$. If we continued this, we would find it to be true that for all column vectors of A that: $\vec{a_i}^T\vec{a_j} = 0$ when $i \neq j$.

In abstract Algebra, a group is a set equipped with a binary operation that combines any two elements to form a third element in such a way that four conditions called group axioms are satisfied, namely closure, associativity, identity and invertibility. We can see with these axioms that orthogonal matrices form a group. We can see that they satisfy the closure axiom since the product of any two orthogonal matrices results in an orthogonal matrix (Proof is left as an exercise to the reader). We can see that it is satisfies the Identity condition since the Identity matrix is itself an orthogonal matrix. From the definition of an orthogonal matrix we know $A^T = A^{-1}$, so it naturally satisfies the inverse axiom. Finally, we know that it satisfies the associativity axiom since $A(BC) = (AB)C$ if A,B,C are all $n \times n$ matrices. Since we know orthogonal matrices have inverses, they must be $n \times n$. Orthogonal matrices have a lot of interesting properties to them, which make them of great interest to mathematicians, myself included.