Orthogonally diagonalizing Symmetric Matrices. If $A = (a_{ij})$ is a (not necessarily square) matrix, the transpose of $A$ denoted $A^T$ is the matrix with $(i, j)$ entry $(a_{ji})$. It is gotten from $A$ by exchanging the $i$th row with the $i$th column, or by “reflecting across the diagonal.” Throughout this note, all matrices will have real entries.

The following are properties satisfied by the transpose.

**Lemma.**

1. $(AB)^T = B^T A^T$
2. $(A^T)^{-1} = (A^{-1})^T$
3. $(A + B)^T = A^T + B^T$

**Definition.** A matrix $A$ is called symmetric if $A = A^T$.

Symmetric matrices have very nice properties. In particular they are orthogonally diagonalizable. This means that if $A$ is symmetric, there is a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ for $\mathbb{R}^n$ consisting of eigenvectors for $A$ so that the vectors in $\mathcal{B}$ are pairwise orthogonal! Another way of saying this is that there exists a matrix $P$ (with real entries) such that $PP^T = P^TP = I$ and $P^TAP$ is a diagonal matrix.

**Definition.** A matrix $P$ such that $P^{-1} = P^T$ is called an orthogonal matrix.

Let $x \cdot y$ denote the usual dot product on $\mathbb{R}^n$. Notice this can be written $x \cdot y = x^T y$, that is ordinary matrix multiplication of the “row vector” $x^T$ and the column matrix $y$.

In particular, if $A$ is symmetric $(Ax) \cdot y = x \cdot (Ay)$.

**Fact.** Symmetric matrices always have real eigenvalues (and hence real eigenvectors).

Moreover,

**Theorem.** If $A$ is symmetric, then eigenvectors of $A$ with distinct eigenvalues are orthogonal.

**Proof** Let $v$ and $w$ be eigenvectors for a symmetric matrix $A$ with different eigenvalues $\lambda_1$ and $\lambda_2$. Then $Av \cdot w = \lambda_1(v \cdot w)$ but also $Av \cdot w = (v \cdot Aw) = \lambda_2(v \cdot w)$, so that $\lambda_1(v \cdot w) = \lambda_2(v \cdot w)$, and since $\lambda_1 \neq \lambda_2$, we must have $(v \cdot w) = 0$.

To understand why a symmetric matrix is orthogonally diagonalizable we must use mathematical induction, so we won’t bother. However, we have an algorithm for finding an orthonormal basis of eigenvectors. Let $A$ be an $n \times n$ symmetric matrix.

1. If $A$ has $n$ distinct eigenvalues, then by the theorem above the corresponding eigenvectors are automatically orthogonal. To get orthonormality, just divide each eigenvector by its length.
2. Suppose $A$ has a repeated eigenvalue $\lambda$. Find a basis (of eigenvectors) $\{v_1, v_2, \ldots, v_k\}$ for $N(A - \lambda I)$. Since $A$ is diagonalizable, there will the same number of eigenvectors corresponding to eigenvalue $\lambda$ as the number of times $\lambda$ appears as a root of the characteristic polynomial of $A$. Apply the Gram-Schmidt process to get an orthogonal basis of eigenvectors $\{x_1, x_2, \ldots, x_k\}$.
3. Repeat the above step for each repeated eigenvalue. Putting all of these bases for $N(A - \lambda_i I)$ together we will have an orthonormal basis.
0.2. **Orthogonal Matrices.** Orthogonal matrices have useful properties as well. For example, if \( v, w \in \mathbb{R}^n \) and we let \( \theta \) be the angle between them, then \( \cos(\theta) = v \cdot w / (||v|| \cdot ||w||) \). Exercise [3] below implies that if \( P \) is orthogonal then
\[
P v \cdot P w / (||P v|| \cdot ||P w||) = \cos(\theta)
\]
so that the linear transformation \( f(x) = P x \) preserves length and preserves the cosine of the angle between any two vectors.

As we observed above, the \( P \) is an orthogonal matrix if and only if its columns form an orthonormal basis for \( \mathbb{R}^n \). Let us figure out all real orthogonal \( 2 \times 2 \) matrices. A matrix \( P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) is orthogonal if \( P P^T = I \) so that
\[
\begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\]
Thus \( a^2 + b^2 = c^2 + d^2 = 1 \) and \( ac + bd = 0 \). The first condition implies that \( (a, b) = (\cos(\theta), \sin(\theta)) \) and \( (c, d) = (\cos(\phi), \sin(\phi)) \) for some angles \( 0 \leq \theta, \phi < 2\pi \). This is because the points \( (a, b) \) and \( (c, d) \) are on a circle of radius 1.

Now the second equation:
\[
0 = ac + bd = \cos(\theta) \cos(\phi) + \sin(\theta) \sin(\phi) =
\frac{\cos(\theta - \phi) + \cos(\theta + \phi)}{2} + \frac{\cos(\theta - \phi) - \cos(\theta + \phi)}{2} = \cos(\theta - \phi)
\]

The equation \( 0 = \cos(x) \) implies that \( x = \pi/2 \) or \( 3\pi/2 \), assuming \( 0 \leq x < 2\pi \), so that \( \theta = \phi \pm \pi/2 \), so \( (c, d) = (\cos(\phi), \sin(\phi)) = \pm (\sin(\theta), \cos(\theta)) \). Putting it all together we have two types of orthogonal \( 2 \times 2 \) matrices:
\[
\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \quad \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}
\]

Notice that \( \det(P) = \pm 1 \), which is a general fact you will prove in the exercises. Since we may choose \( \theta \) to be any angle in \([0, 2\pi)\) so there are infinitely many \( 2 \times 2 \) orthogonal matrices. The linear transformation \( f(x) = P x \) rotates the vector \( x \) through an angle of \( \theta \) or \( -\theta \) depending on \( \det(P) \). Another way of seeing that the above are the only possible \( 2 \times 2 \) orthogonal matrices is to observe that, in \( \mathbb{R}^2 \), for any fixed vector \( x \) with \( ||x|| = 1 \) there are exactly two vectors \( y \) with \( ||y|| = 1 \) and \( x \cdot y = 0 \).

Challenge problem: Define \( Q(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \) and \( P(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \).
\( \text{(Notice that } \det(P(\theta)) = 1 \text{ while } \det(Q(\theta)) = -1. \) Show that \( P(\theta) P(\phi) = P(\theta + \phi) \), \( Q(\theta) Q(\phi) = Q(\phi - \theta) \) and \( P(\theta) Q(\phi) = Q(\phi - \theta) \).

The \( n \times n \) orthogonal matrices of determinant +1 can also be visualized as “rigid motions” in space. In \( \mathbb{R}^3 \) the effect of multiplying a vector \( x \) by such an orthogonal matrix is to rotate \( x \) through two angles in succession. Notice that the matrix that switches the \( x \)-axis and the \( y \)-axis while fixing the \( z \)-axis is not a rigid motion, but has determinant \(-1\).

**Exercise (1).** Verify that \( (A^T)^{-1} = (A^{-1})^T \), using the fact that \( (AB)^T = B^T A^T \).
Exercise (2). Show that for a square matrix $A$, $(Ax) \cdot y = x \cdot (A^T y)$.

Exercise (3). Assume that $P$ is orthogonal. Show that $v \cdot w = (Pv) \cdot (Pw)$.

Exercise (4). Suppose that $\{v_1, \ldots, v_k\}$ is an orthogonal set in $\mathbb{R}^n$. Show that the set $\{Pv_1, \ldots, Pv_k\}$ is also orthogonal if $P$ is an orthogonal matrix.

Exercise (5). Assume that $P$ is orthogonal and symmetric. Show that $P^2 = I$.

Exercise (6). Show that if $P$ and $Q$ are orthogonal matrices, then so is $PQ$. Show that $\det(P) = \pm 1$.

Exercise (7). Show that the set, $S$ of symmetric matrices is a subspace of $M_{n,n}$. Determine $\dim(S)$.

Exercise (8). Let $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Find an orthogonal matrix $P$ so that $PBP^{-1}$ is diagonal.