• Symmetric matrices.

So far we have exclusively studied $2 \times 2$ matrices. Obviously, nothing stops us from studying their larger-size counterparts ($n \times n$ matrices). Going from $2 \times 2$ to $n \times n$ is not completely vis-a-vis: Basic properties such as associativity, distributivity, the identity matrix, etc. are entirely parallel, whereas we encounter ramifications which didn’t arise in the $2 \times 2$ case. In order to deal with such ramifications, we cope with the notion of ‘vector spaces’. So far we were able to do away with ‘vector spaces’, but that was only to the extent possible. Already in the first course in linear algebra (Math 290), you (most of you) have been exposed to some abridged version of the theory of ‘vector spaces’. To anyone who sees it for the first time, it looks pretentious, full of abstruse formalism. Contrary to such uninviting impression, the notion of ‘vector spaces’ is indispensable in mathematics, as in the entire math depends on it. It is probably not reasonable (as in too much) to expect that your professor can make a strong case that that is true, in a one-semester course. But in this class I plan to present some supporting evidence of it within the framework of linear algebra. One major advantage of working with the notion of vector spaces is that matrices are naturally recalibrated as ‘linear transforms’. Eigenvalues make sense for linear transforms. This seemingly minor tweak of the pre-existing notion turns out to be a game-changer.

For an $n \times n$ matrix $A$, it is not always true that two eigenvectors associated with one eigenvalue of a matrix $A$ are scalar multiples of each other. Therefore, in order to form $Q$ to diagonalize $A$ (provided it is feasible) you need to choose a ‘basis’ (one key concept associated with the notion of vector spaces) of the so-called ‘eigenspace’ of $A$. At this point, without adequate understanding of the notion of vector spaces, you would not be able to move forward.

I said on Day 1 that I shoot for ‘spectral theory’. Today I am going to cover the archetype of spectral theory, namely, the $2 \times 2$ case of it. We prove that a $2 \times 2$ symmetric matrix $A$ with real number entires is diagonalizable by a real orthogonal matrix $Q$. The proof the same statement for the $n \times n$ case is abstract, and you really need to rely on the notion of abstract vector spaces over ‘the complex number field’. So, today is the last day we stick with $2 \times 2$ case. Let’s start with:
Definition (Transpose).

For a matrix \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \), define its transpose as \( A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \).

So, the transpose of \( A \) means you just swap the top-right and the bottom-left entries of \( A \).

Repeat: If \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \).

Formula. For \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) and \( B = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \),

(1) \( (A^T)^T = A \).
(2) \( (A + B)^T = A^T + B^T \).
(3) If \( t \) is a scalar, then \( (tA)^T = tA^T \).
(4) \( (AB)^T = B^T A^T \).

Exercise 1. Prove part (4) of the above formula.

Definition (Symmetric matrix).

Suppose \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) satisfies

\( A^T = A \),

then we say \( A \) is a symmetric matrix.
• Clearly, \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) is a symmetric matrix precisely when the top-right entry and the bottom-left entry of \( A \) are equal: \( b = c \).

• In other words, a general form of a 2 \( \times \) 2 symmetric matrix is

\[
A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.
\]

**Example 1.** \[
\begin{bmatrix} 3 & 5 \\ 5 & -1 \end{bmatrix}
\]
is a symmetric matrix.

**Example 2.** \[
\begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix}
\]
is a symmetric matrix.

**Example 3.** \[
\begin{bmatrix} 7 & 7 \\ 7 & 7 \end{bmatrix}
\]
is a symmetric matrix.

**Example 4.** \[
\begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}
\]
is a symmetric matrix.

**Example 5.** \[
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]
is a symmetric matrix.

**Example 6.** \[
\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\]
is a symmetric matrix.

**Formula 2.** Let \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) be a matrix, not necessarily symmetric.

Then

1. \( A + A^T \) is symmetric.
2. \( AA^T \) is symmetric.
3. \( A^T A \) is symmetric.
Example 7. \( A = \begin{bmatrix} 2 & 4 \\ 3 & 0 \end{bmatrix} \) is not a symmetric matrix. Its transpose is
\[
A^T = \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix}.
\]

This \( A^T \) is not a symmetric matrix either. However,

(1) \[
A + A^T = \begin{bmatrix} 2 & 4 \\ 3 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 7 & 0 \end{bmatrix}
\]
is a symmetric matrix.

(2) \[
A A^T = \begin{bmatrix} 2 & 4 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 20 & 6 \\ 6 & 9 \end{bmatrix}
\]
is a symmetric matrix.

(3) \[
A^T A = \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 8 \\ 8 & 16 \end{bmatrix}
\]
is a symmetric matrix.

Note. As this example shows, in general \( AA^T \neq A^T A \).

The next subject is related to the last lecture (“Review of Lectures – IX”):

**Orthogonal matrices revisited.**

We have already previously defined orthogonal matrices, in “Review of Lectures – IX”. Let’s recall:

**Definition 1 (Orthogonal matrices).** Matrices of the forms
\[
A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix},
\]
(\( \theta \): a real number) are called orthogonal matrices.
There is actually another, equivalent, definition of orthogonal matrices.

\textbf{Definition 2 (Alternative definition of orthogonal matrices).}

\textbf{Suppose } \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ satisfies } AA^T = I,

\text{then we say } A \text{ is an orthogonal matrix.}

The above “alternative definition” (Definition 2) seemingly differs from our original definition (Definition 1 in the previous page). However, it turns out that those two definitions are mathematically equivalent:

\[
\begin{array}{c}
\text{Definition 1 } \iff \text{ Definition 2.}
\end{array}
\]

\textbf{Proof of equivalence between the two definitions (of orthogonal matrices).}

We need to prove

(i) \ [\text{Definition 1 } \implies \text{ Definition 2}] \quad \text{ and }

(ii) \ [\text{Definition 2 } \implies \text{ Definition 1}].

- First, let’s prove (i). This part is easier of the two. For

\[
A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix},
\]

we have
\[ AA^T = \begin{bmatrix} \cos \theta & - \sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ - \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} (\cos \theta)(\cos \theta) + (-\sin \theta)(-\sin \theta) & (\cos \theta)(\sin \theta) + (-\sin \theta)(\cos \theta) \\ (\sin \theta)(\cos \theta) + (\cos \theta)(-\sin \theta) & (\sin \theta)(\sin \theta) + (\cos \theta)(\cos \theta) \end{bmatrix} = \begin{bmatrix} (\cos \theta)^2 + (\sin \theta)^2 & 0 \\ 0 & (\sin \theta)^2 + (\cos \theta)^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \]

Next, for
\[ B = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}, \]
we have
\[ BB^T = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix} = \begin{bmatrix} (\cos \theta)(\cos \theta) + (\sin \theta)(\sin \theta) & (\cos \theta)(\sin \theta) + (\sin \theta)(-\cos \theta) \\ (\sin \theta)(\cos \theta) + (-\cos \theta)(\sin \theta) & (\sin \theta)(\sin \theta) + (-\cos \theta)(-\cos \theta) \end{bmatrix} = \begin{bmatrix} (\cos \theta)^2 + (\sin \theta)^2 & 0 \\ 0 & (\sin \theta)^2 + (\cos \theta)^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \]

Here, in the above, we relied on
\[ \boxed{(\cos \theta)^2 + (\sin \theta)^2 = 1}. \]
Second, let’s prove (ii). Suppose \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) satisfies \( AA^T = I \), then

\[
\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]

That is

\[
\begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]

So

\[
\begin{cases}
  a^2 + b^2 = 1, \\
  c^2 + d^2 = 1, \\
  ac + bd = 0.
\end{cases}
\]

Now we rely on the following, from “Review of Lecutres – IX”:

**Fact.** Let \( a \) and \( b \) be real numbers. Suppose they satisfy

\[a^2 + b^2 = 1.\]

Then there exists a real number \( \theta \) such that

\[a = \cos \theta, \quad b = \sin \theta.\]

If you use this, you conclude

\[a = \cos \theta, \quad b = \sin \theta\]

from \( a^2 + b^2 = 1 \) (in (\( * \))), and also conclude

\[c = \cos \phi, \quad d = \sin \phi.\]

from \( c^2 + d^2 = 1 \) (in (\( * \))).
Now substitute these into the equation
\[ ac + bd = 0 \]
(in \((*)\)):
\[
\left( \cos \theta \right) \left( \cos \phi \right) + \left( \sin \theta \right) \left( \sin \phi \right) = 0.
\]
In other words
\[
\cos \left( \theta - \phi \right) = 0.
\]
(Here we used one of the ‘trig identities’. See below.) So
\[
\theta - \phi = \frac{\pi}{2} + k \pi,
\]
that is,
\[
\phi = \theta - \frac{\pi}{2} - k \pi,
\]
with some integer \( k \). Then it follows that
\[
c = \cos \phi
\]
\[= \cos \left( \theta - \frac{\pi}{2} - k \pi \right)\]
\[= \pm \sin \theta\]
\((\text{double sign is ‘+’ when } k \text{ is even, ‘−’ when } k \text{ is odd})\), and
\[
d = \sin \phi
\]
\[= \sin \left( \theta - \frac{\pi}{2} - k \pi \right)\]
\[= \mp \cos \theta\]
\((\text{double sign is ‘−’ when } k \text{ is even, ‘+’ when } k \text{ is odd})\). So we conclude
\[
A = \begin{bmatrix} a & b \\ \pm b & \mp a \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ \pm \sin \theta & \mp \cos \theta \end{bmatrix} \quad \text{(double sign in the same order)}. \quad \Box
\]
• In the above proof, we have used one of the following:

**Formula.**

\[
\begin{align*}
\cos(\theta + \phi) &= \left( \cos \theta \right) \left( \cos \phi \right) - \left( \sin \theta \right) \left( \sin \phi \right), \\
\sin(\theta + \phi) &= \left( \sin \theta \right) \left( \cos \phi \right) + \left( \cos \theta \right) \left( \sin \phi \right), \\
\cos(\theta - \phi) &= \left( \cos \theta \right) \left( \cos \phi \right) + \left( \sin \theta \right) \left( \sin \phi \right), \\
\sin(\theta - \phi) &= \left( \sin \theta \right) \left( \cos \phi \right) - \left( \cos \theta \right) \left( \sin \phi \right).
\end{align*}
\]

• **Spectral theory – I (Archetype).**

Now we are ready to talk about ‘spectral theory’. Here is the main statement:

**Spectral Theorem (2 \times 2 case).** Let \( A \) be a 2 \times 2 symmetric matrix:

\[
A = \begin{bmatrix}
a & b \\
b & d
\end{bmatrix}.
\]

Suppose \( a, b, c \) and \( d \) are all real numbers. Then there exists an orthogonal matrix \( Q = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \), where \( p, q, r \) and \( s \) are all real numbers, such that

\[
Q^{-1}AQ = \begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix},
\]

where \( \lambda_1 \) and \( \lambda_2 \) are the eigenvalues of \( A \) (they may or may not coincide).
Proof of Spectral theorem.

First, with a suitable scalar $t$, we can make

$$A - tI = \begin{bmatrix} a - t & b \\ b & d - t \end{bmatrix}$$

satisfy $(a - t) + (d - t) = 0$. Let’s call this new matrix as $B$:

$$B = A - tI = \begin{bmatrix} a' & b \\ b & d' \end{bmatrix}$$

where $a' + d' = 0$. Next, with a suitable scalar $s$, we can make

$$sB = \begin{bmatrix} sa' & sb \\ sb & sd' \end{bmatrix}$$

satisfy

$$sa' = \cos \theta, \quad sb = \sin \theta \quad \text{(for some real number } \theta) \).$$

(Indeed, $s = \left( \sqrt{a'^2 + b^2} \right)^{-1}$.) Then from $a' + d' = 0$ we also have

$$sd' = - \cos \theta$$

for the same $\theta$. Let’s call this new matrix as $C$:

$$C = sB = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & - \cos \theta \end{bmatrix}.$$ 

We already know that $C$ has eigenvalues $\pm 1$, and moreover

$$x_\pm = \begin{bmatrix} \sin \theta \\ - \left( \cos \theta \right) \pm 1 \end{bmatrix}$$

is an eigenvector of $C$ associated with $\lambda = \pm 1$ (the double sign in the same order).
Here, recall that for a vector \( \mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix} \), we define its norm as
\[
\| \mathbf{u} \| = \sqrt{u^2 + v^2}.
\]

Also recall that, if \( \mathbf{u} \neq \mathbf{0} \), then
\[
\left\| \frac{1}{\| \mathbf{u} \|} \mathbf{u} \right\| = 1.
\]

Now, \( \| \mathbf{x}_\pm \| = \sqrt{2 \mp 2 \cos \theta} \). Accordingly, set
\[
\mathbf{y}_\pm = \frac{1}{\| \mathbf{x}_\pm \|} \mathbf{x}_\pm = \frac{1}{\sqrt{2 \mp 2 \cos \theta}} \begin{bmatrix} \sin \theta \\ -\left( \cos \theta \right) \pm 1 \end{bmatrix}
\]
(where the double sign in the same order is kept intact). Then \( \mathbf{y}_\pm \) is another eigenvector of \( \mathbf{C} \) associated with \( \lambda = \pm 1 \). Indeed, \( \mathbf{y}_\pm \) is a non-zero scalar multiple of \( \mathbf{x}_\pm \), and \( \mathbf{x}_\pm \) is an eigenvector of \( \mathbf{C} \) associated with \( \lambda = \pm 1 \). Now, set
\[
\mathbf{Q} = \begin{bmatrix} \mathbf{y}_+ & \mathbf{y}_- \end{bmatrix} = \begin{bmatrix} \frac{\sin \theta}{\sqrt{2 - 2 \cos \theta}} & \frac{\sin \theta}{\sqrt{2 + 2 \cos \theta}} \\ -\left( \cos \theta \right) + 1 & -\left( \cos \theta \right) - 1 \end{bmatrix}.
\]

Then \( \mathbf{Q} \) is an orthogonal matrix. Indeed, by direct calculation \( \mathbf{Q}^T \mathbf{Q} \) turns out to equal \( \mathbf{I} \), and hence \( \mathbf{Q} \mathbf{Q}^T \) also equals \( \mathbf{I} \). Now, with this \( \mathbf{Q} \) we can make \( \mathbf{Q}^{-1} \mathbf{C} \mathbf{Q} \) diagonal, since the columns of \( \mathbf{Q} \) are eigenvectors of \( \mathbf{C} \) associated with two distinct eigenvalues of \( \mathbf{C} \). Then with the same \( \mathbf{Q} \) we can make \( \mathbf{Q}^{-1} \mathbf{B} \mathbf{Q} \) diagonal, since \( \mathbf{B} \) is a scalar multiple of \( \mathbf{C} \). Then in turn we can make \( \mathbf{Q}^{-1} \mathbf{A} \mathbf{Q} \) diagonal, since \( \mathbf{A} \) is \( \mathbf{B} \) plus a scalar times \( \mathbf{I} \). \( \square \)

**Exercise 2.** (1) For \( \mathbf{x}_\pm = \begin{bmatrix} \sin \theta \\ -\left( \cos \theta \right) \pm 1 \end{bmatrix} \), verify
\[
\| \mathbf{x}_\pm \| = \sqrt{2 \mp 2 \cos \theta} \quad \text{(double sign in the same order)}.
\]
(2) Perform the actual calculation $Q^TQ$ for

$$Q = \begin{bmatrix} y_+ & y_- \end{bmatrix} = \begin{bmatrix} \frac{\sin \theta}{\sqrt{2 - 2 \cos \theta}} & \frac{\sin \theta}{\sqrt{2 + 2 \cos \theta}} \\ \frac{-\left( \cos \theta \right) + 1}{\sqrt{2 - 2 \cos \theta}} & \frac{-\left( \cos \theta \right) - 1}{\sqrt{2 + 2 \cos \theta}} \end{bmatrix},$$

and verify that it equals $I$.

**Exercise 3.** Diagonalize by an orthogonal matrix:

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}.$$