Where do eigenvalues/eigenvectors/eigenfunctions come from, and why are they important anyway?

I. **Background** (from Ordinary Differential Equations)

Consider the simplest example of a harmonic oscillator (think of a vibrating string)

\[
\frac{d^2}{dx^2} u(x) + \lambda^2 u(x) = 0
\]

We know this equation has the linearly independent solutions

\[
\begin{align*}
    u_1(x) &= \sin(\lambda x) \\
    u_2(x) &= \cos(\lambda x)
\end{align*}
\]

where \( \lambda \) is related to the frequency of oscillation. As \( \lambda \) increases, the frequency of oscillation increases.

We can gain more insight by imposing boundary conditions \( u(0) = u(L) = 0 \). The homogeneous solutions take the form of

\[
\begin{align*}
    u(x) &= Au_1(\lambda x) + Bu_2(\lambda x) \\
          &= A \sin(\lambda x) + B \cos(\lambda x)
\end{align*}
\]

Applying the first boundary condition, \( u(0) = 0 \), we see that

\[
0 = u(0) = B
\]

Using the second boundary condition, we conclude that
\[ 0 = u(L) = A \sin(\lambda L) \]

Which means that \( \lambda L = n\pi \) for integer values of \( n \), \( n = 1, 2, 3, \ldots \). This gives us an infinite sequence of simple discrete values

\[ \lambda_n = \frac{n\pi}{L} \]

Writing the simple harmonic oscillator equation as

\[ D^2[u] = \frac{d^2}{dx^2} u(x) = -\lambda^2 u(x) \]

we see that it is of the form

\[ Lu = \beta u \]

where \( \beta \) is a constant, a “special” value, called an eigenvalue. (The word “eigen” means characteristic, or singular, or unique in German.)

II. Linear Algebra Application

In \( \mathbb{R}^n \), the inner product of two column vectors is given by

\[ u \cdot v = u^T v = (v_1, v_2, \ldots, v_n) \cdot (u_1, u_2, \ldots, u_n) = u_1v_1 + u_2v_2 + \ldots + u_nv_n \]

For symmetric matrices, \( A = A^T \), we can say something more – the vectors (eigenvectors) associated with distinct eigenvalues are orthogonal.

\[ Au_1 = \lambda_1 u_1 \]
\[ Au_2 = \lambda_2 u_2 \]

Multiplying the first equation by \( u_2^T \) and the second by \( u_1^T \) we get

\[ u_2^T Au_1 = u_2^T \lambda_1 u_1 = \lambda_1 u_2^T u_1 \]
\[ u_1^T Au_2 = u_1^T \lambda_2 u_2 = \lambda_2 u_1^T u_2 \]

Since all of the expressions are scalars, we have

\[ u_1^T Au_2 = (u_1^T Au_2)^T = u_2^T A^T u_1 \]

Subtracting the two equations we have the identity
If $A$ is symmetric, that is if $A = A^T$, and if $\lambda_1 \neq \lambda_2$ then we must have $u_1^T u_2 = 0$. This leads to the following results

**Theorem:** If $A$ is symmetric ($A = A^T$) then eigenvectors corresponding to different eigenvalues are orthogonal.

**Corollary:** If an $n \times n$ symmetric matrix has distinct eigenvalues, then it has $n$ linearly independent (and orthogonal) eigenvectors. This set of eigenvectors forms a basis.

### III. Bases of eigenvectors

If a matrix $A$ has a complete set of eigenvectors that can be used as a basis, then solving a linear system $Au = f$ becomes very simple.

Let $\{e_1, e_2, ..., e_n\}$ be a basis of eigenvectors, which have been normalized with length 1. Given a right hand side $f \in \mathbb{R}^n$, we can expand $f$ in terms of coordinates

$$f = \sum_{k=1}^{n} f_k e_k$$

The coordinates, $f_k$, are easy to compute, since by orthogonality

$$e_j^T f = \sum_{k=1}^{n} f_k e_j^T e_k = f_j$$

Similarly, the solution has a basis expansion

$$u = \sum_{k=1}^{n} u_k e_k$$

where the basis coefficients are not known. Substituting these into $Au = f$, we get

$$\sum_{k=1}^{n} f_k e_k = f = \sum_{k=1}^{n} u_k A e_k = \sum_{k=1}^{n} u_k \lambda_k e_k$$

Since the basis coefficients (coordinates) are unique, we must have

$$f_k = \lambda_k u_k,$$

$$u_k = \frac{f_k}{\lambda_k}.$$
So, the solution of the linear system amounts to division!

\[ \sum_{k=1}^{n} f_k e_k = f \rightarrow u = \sum_{k=1}^{n} u_k e_k = \sum_{k=1}^{n} \frac{f_k}{\lambda_k} e_k \]

IV. Geometric Interpretation of Eigenvalues

If you look at the image of the unit ball, \( \|u\| = 1 \), under the mapping \( A \), then the largest eigenvalue is equal to the maximum “stretching” of \( A \).

\[ \lambda_{\text{max}} = \max \|Au\| \]

The eigenvector is the direction of maximum stretching.

V. Calculation of Eigenvalues and Eigenvectors for a finite dimensional matrix

From the definition of an eigenvalue and an eigenvector, we have

\[ Ax = \lambda x = \lambda Ix \]

which leads to \( 0 = Ax - \lambda Ix = (A - \lambda I)x \). This can have non-zero solutions only if the matrix is singular (not invertible). Therefore, we must have \( \det(A - \lambda I) = 0 \). This results in a polynomial equation of order \( n \) for an \( nxn \) matrix.

VI. Diagonalization of Linear Operators

Another way to view the behavior of eigenvalues is the process of diagonalization. If we consider the differentiation operator, \( D = \frac{d}{dx} \), we see that

\[ D[e^{ikx}] = ike^{ikx} \]

So the function \( e^{ikx} \) is an eigenfunction, with eigenvalue \( \lambda_k = ik \). The eigenfunctions are orthogonal with respect to the inner product

\[ \langle e^{imx}, e^{inx} \rangle = \frac{1}{2\pi} \int_{0}^{2\pi} e^{imx} e^{-inx} dx = \frac{1}{2\pi} \int_{0}^{2\pi} e^{i(m-n)x} dx = \]

\[ \frac{1}{2\pi} \left( \cos((m-n)x) + i \sin((m-n)x) \right)_{x=2\pi} = \begin{cases} 1, m = n \\ 0, m \neq n \end{cases} \]

We can calculate the coordinates (i.e. Fourier coefficients) by means of the formula

\[ \hat{u}_k = \langle u(x), e^{ikx} \rangle = \frac{1}{2\pi} \int_{0}^{2\pi} u(x)e^{-ikx} dx \]
With respect to this basis, the differential operator \( D \) is diagonal

\[
\begin{bmatrix}
1 & e^{ix} & e^{ix} & ... & e^{inx} \\
e^{ix} & ie^{ix} & 2ie^{ix} & ... & nie^{inx} \\
e^{2ix} & 0 & 0 & 2i & ... \\
... & ... & ... & ... & ...
\end{bmatrix}
= \begin{bmatrix}
0 & 0 & 0 & ... & 0 \\
0 & i & 0 & ... & 0 \\
0 & 0 & 2i & ... & 0 \\
0 & 0 & 0 & ... & ni \\
0 & ie^{ix} & 2ie^{ix} & ... & nie^{inx}
\end{bmatrix}
\]

Since the upper left coefficient is 0, the matrix is singular. We know that \( D \) is not 1-1, since \( D[1]=0 \) and \( 1 \neq 0 \).

VII. Filtering

If you look at the Fourier coefficients of a continuous function, they generally will look like the following

The coefficients will gradually decay. Large values of \( k \) are associated with higher frequency components, usually with “noise.” Eliminating these components (“low pass filtering”) generally leads to a smoother function.

VIII. Eigenfunctions of a Linear Operator

If we have a linear operator, \( L \), with a complete set of orthonormal eigenfunctions, and an inhomogeneous equation
\[ Lu = f \]

we can solve this problem using a basis of eigenfunctions.

We can expand \( f \) in terms of the basis \( f(x) = \sum_k f_k \phi_k(x) \) (with known coefficients, given by \( f_k = \langle \phi_k, f \rangle \) and the solution given as \( u(x) = \sum_k u_k \phi_k(x) \). Substituting this into the inhomogeneous equation, we get

\[
\sum_k f_k \phi_k(x) = f(x) = Lu = L\left(\sum_k u_k \phi_k(x)\right) = \sum_k u_k \lambda_k \phi_k(x)
\]

Equating the coefficients of \( \phi_k(x) \), we get the relation

\[
u_k = \frac{f_k}{\lambda_k}\]

Since \( \lambda_k \to \infty \), and \( |f_k| \) are bounded (in fact they go to zero also), we must have

\[
u_k \to 0
\]

The more quickly the coefficients go to zero, the smoother (more differentiable) the function.

**IX. Summary**

- If an nxn matrix is symmetric and has distinct eigenvalues, then it has complete set of n eigenvectors, which may be used as an orthonormal basis.
- Eigenvectors and eigenfunctions are most often identified with fundamental modes of vibration or oscillation.
- Eigenvalues are associated with the frequencies of vibration or oscillation.
- The amplitude of the oscillation goes down asymptotically as the frequency increases. The faster the rate of the decay, the smoother the function.
- “Noise” is generally associated with high frequency components. Cutting off high frequency components (setting the Fourier coefficients to zero) results in a smoother function.