Eigenvectors and eigenvalues

COMS 3251 Fall 2022 (Daniel Hsu)

1 Eigenvectors and eigenvalues

Let us consider an $n \times n$ matrix A as a linear operator on $\mathbb{R}^n$: it linearly transforms n-vectors into other n-vectors. One way we might try to understand A (for some application) is by decomposing each operator into "simpler" parts, each of which (we hope) is easier to understand.

For example, suppose we would like to understand the behavior of the linear operator in the context of a *linear dynamical system*:

$$\mathbf{x}_{t+1} \;=\; A\mathbf{x}_t \quad \text{for } t \in \{0, 1, 2, \ldots\}.$$

Here, the vector $\mathbf{x}_t$ is regarded as the state of a system at time t; the state of the system evolves over (discrete) time by applying the linear operator A to the state. Evidently, the state at time t is $A^t\mathbf{x}_0$, where A^t is just shorthand for A written out t times. How can we understand the long term behavior of the system as a function of the initial state $\mathbf{x}_0$?

One can, of course, simply simulate the evolution of the system starting at $\mathbf{x}_0$ by repeatedly multiplying $\mathbf{x}_0$ by A. But this is expensive for large t and not possible in finite time for $t \to \infty$. Is there a better way?

Example. Let A be the 2×2 matrix given as follows:

$$A \;=\; \begin{bmatrix} -1/2 & 5/2 \\ 0 & 2 \end{bmatrix}.$$

- Let $\mathbf{v} = (1, 1)$. Then $A\mathbf{v} = (2, 2) = 2\mathbf{v}$. In other words, the effect of the operator A on $\mathbf{v}$ is simply to scale $\mathbf{v}$ by 2. So $A^t\mathbf{v} = 2^t\mathbf{v}$ for all t. Vectors on the line $\{c\mathbf{v} : c \in \mathbb{R}\}$ stay there, but move away from $\mathbf{0}$.

- Let $\mathbf{w} = (1, 0)$. Then $A\mathbf{w} = (-1/2, 0) = -(1/2)\mathbf{w}$. In other words, the effect of the operator A on $\mathbf{w}$ is simply to scale $\mathbf{w}$ by $-1/2$. So $A^t\mathbf{w} = (-1/2)^t\mathbf{w}$ for all t. Vectors on the line $\{c\mathbf{w} : c \in \mathbb{R}\}$ stay there, but move closer to $\mathbf{0}$ on the opposite side.

- Note that $\{\mathbf{v}, \mathbf{w}\}$ is linearly independent, so every vector in $\mathbb{R}^2$ can be written as a linear combination of $\mathbf{v}$ and $\mathbf{w}$. So if $\mathbf{x} = a\mathbf{v} + b\mathbf{w}$, then

$$A^t\mathbf{x} \;=\; A^t(a\,\mathbf{v} + b\,\mathbf{w}) \;=\; a\,A^t\mathbf{v} + b\,A^t\mathbf{w} \;=\; a\,2^t\,\mathbf{v} + b\left(-\frac{1}{2}\right)^t\mathbf{w}. \quad \blacksquare$$

The vectors $\mathbf{v}$ and $\mathbf{w}$ in the previous example are very special to the matrix A: they are eigenvectors of the matrix A. We say a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ is an *eigenvector* of an $n \times n$ matrix A if there is a scalar λ such that $A\mathbf{v} = \lambda\mathbf{v}$; this scalar λ is called the *eigenvalue* corresponding to the eigenvector $\mathbf{v}$. We also say that a scalar λ is an eigenvalue of an $n \times n$ matrix A if there exists a non-zero vector $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. Note that if $\mathbf{v}$ is an eigenvector of A, then every non-zero vector in $\operatorname{span}(\{\mathbf{v}\})$ is also an eigenvector of A with the same eigenvalue.

Given a scalar λ, it is easy to check whether λ is an eigenvalue of A. Consider the following equivalences:

$$A\mathbf{v} \;=\; \lambda\mathbf{v} \quad\Leftrightarrow\quad A\mathbf{v} \;=\; \lambda I\mathbf{v} \quad\Leftrightarrow\quad A\mathbf{v}-\lambda I\mathbf{v} \;=\; \mathbf{0} \quad\Leftrightarrow\quad (A-\lambda I)\mathbf{v} \;=\; \mathbf{0}.$$

So, $\mathbf{v}$ is an eigenvector of A with corresponding eigenvalue λ if and only if $\mathbf{v}$ is a non-zero vector in the nullspace of $A - \lambda I$. And we have seen previously how to find a basis for the nullspace of any matrix (e.g., via Elimination).

Example. Let A be the matrix defined by

$$A \;=\; \begin{bmatrix} 4 & 2 & 2 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

Suppose a little bird suggests to you that $\lambda = 2$ might be an eigenvalue of A. To check, we "solve for" the nullspace of $A - \lambda I$ for $\lambda = 2$. We transform

$$A - 2I \;=\; \begin{bmatrix} 4 & 2 & 2 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \;=\; \begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

to the following RREF matrix via elementary row operations: $R = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ (after dropping all-zeros rows). The homogeneous system of linear equations

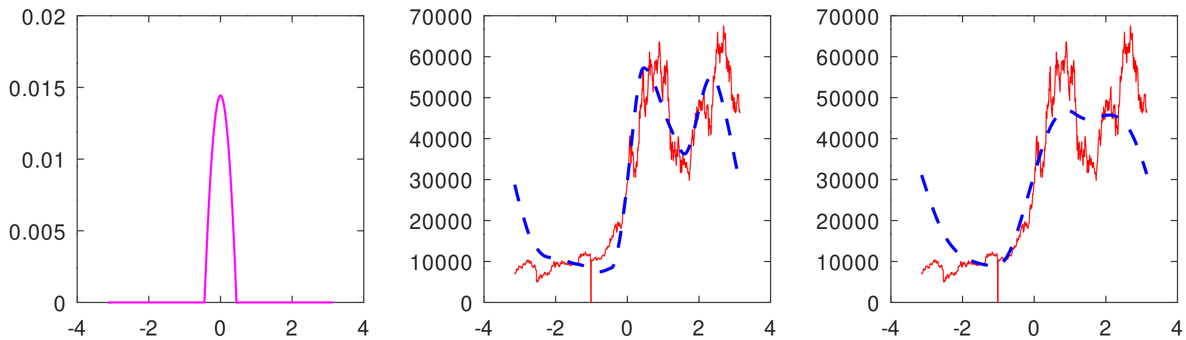

Figure 1: Plots of functions in $\mathsf{C}_{\mathsf{periodic}}([-\pi, \pi], \mathbb{R})$; the horizontal axis corresponds to $t \in [-\pi, \pi]$. Left: a "nice" kernel function $h(t)$. Middle: a "rough" function $f(t)$ (red solid line) and its convolution with $h(t)$ (blue dashed line). Right: same as middle plot but using a wider kernel function.

$(A - 2I)\mathbf{x} = \mathbf{0}$ therefore has 2 free variables, $\mathsf{FV} = \{2, 3\}$, so the nullspace of $A - 2I$ has dimension 2. The special solutions are $\mathbf{s}_2 = (-1, 1, 0)$ and $\mathbf{s}_3 = (-1, 0, 1)$. So, the eigenvectors with eigenvalue 2 are the non-zero vectors in $\operatorname{span}(\{(-1, 1, 0), (-1, 0, 1)\})$. ∎

Important example. The concept of eigenvectors and eigenvalues extends to linear operators on general vector spaces. Consider the inner product space $\mathbb{W} = \mathsf{C}_{\mathsf{periodic}}([-\pi, \pi], \mathbb{R})$ of continuous, real-valued functions on $[-\pi, \pi]$ that are periodic with period 2π, equipped with the inner product

$$\langle f, g \rangle_{\mathbb{W}} \; = \; \int_{-\pi}^{\pi} f(t) g(t) \, \mathrm{d}t.$$

The *convolution operator* $T_h \colon \mathbb{W} \to \mathbb{W}$, based on a specific "kernel" function $h \in \mathbb{W}$, is the operator that transforms a function $f \in \mathbb{W}$ into another function $g = T_h(f)$ by "convolving" f with h:

$$g(t) \; = \; \int_{-\pi}^{\pi} f(t - x) h(x) \, \mathrm{d}x.$$

Convolution operators are important in many applications. For instance, it is common to transform a "rough" function into a "smoother" one by convolving the rough function with a "nice" kernel function. See Figure 1.

We'll show that the functions $e_0(t) = 1$, $e_k(t) = \cos(kt)$, and $f_k(t) = \sin(kt)$ for all $k \in \mathbb{N}$ are eigenvectors of the convolution operator T_h. For simplicity, we assume h is an even function (i.e., $h(x) = h(-x)$). First, we consider the constant function e_0:

$$(T_h e_0)(t) \;=\; \int_{-\pi}^{\pi} h(x)\,\mathrm{d}x \;=\; \langle h, 1\rangle_{\mathbb{W}}\, e_0.$$

So e_0 is an eigenvector of T_h with corresponding eigenvalue $\langle h, e_0\rangle_{\mathbb{W}}$. Next, we consider e_k for $k \in \mathbb{N}$:

$$\begin{aligned}
(T_h e_k)(t) \;&=\; \int_{-\pi}^{\pi} \cos(k(t-x))h(x)\,\mathrm{d}x \\
&=\; \int_{-\pi}^{\pi} (\cos(kt)\cos(kx) + \sin(kt)\sin(kx))h(x)\,\mathrm{d}x \\
&=\; e_k(t)\int_{-\pi}^{\pi} e_k(x)h(x)\,\mathrm{d}x + \sin(kt)\cancelto{0}{\int_{-\pi}^{\pi} \sin(kx)h(x)\,\mathrm{d}x} \\
&=\; \langle e_k, h\rangle_{\mathbb{W}}\, e_k(t),
\end{aligned}$$

where the second equality uses a standard trigonometric identity, and the cancellation comes from the fact that $\sin(kx)h(x)$ is an odd function. So e_k is an eigenvector of T_h with corresponding eigenvalue $\langle e_k, h\rangle_{\mathbb{W}}$. Finally, we consider f_k for $k \in \mathbb{N}$. Using a similar calculation, we find

$$(T_h f_k)(t) \;=\; \langle e_k, h\rangle_{\mathbb{W}}\, f_k(t).$$

So f_k is an eigenvector of T_h with corresponding eigenvalue $\langle e_k, h\rangle_{\mathbb{W}}$.

If instead h is an odd function (i.e., $h(-x) = -h(x)$), we would reach a similar conclusion but with different eigenvalues. It is worth pointing out that every function h is the sum of an even function h_{even} and an odd function h_{odd}. For such a function $h = h_{\mathsf{even}} + h_{\mathsf{odd}}$, we have $T_h = T_{h_{\mathsf{even}}} + T_{h_{\mathsf{odd}}}$. ∎

It was fortunate in our first example that the $n \times n$ matrix A (for $n = 2$) had n linearly independent eigenvectors. This is because they form a basis for $\mathbb{R}^n$, which means we are able to write every vector in $\mathbb{R}^n$ as a linear combination of the eigenvectors of A. It is in this sense that we are able to decompose A into simpler "parts" that are easier to understand. (We make this more precise later in Section 4.)

However, not all $n \times n$ matrices have n linearly independent eigenvectors, as the next examples show. (Also see Appendix B.)

Examples.

- For $\theta \in [0, 2\pi)$, let R be the 2×2 matrix given by

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

  If $\mathbf{v}$ is a non-zero 2-vector, then the matrix-vector product $R\mathbf{v}$ results in a rotation of $\mathbf{v}$ by θ radians. This is on the same line as $\mathbf{v}$ if and only if $\theta = 0$ (no rotation) or $\theta = \pi$ (rotate 180°). So if $\theta \notin \{0, \pi\}$, then R does not have any eigenvectors.

- For a non-zero linear functional $\mathbf{w}^\intercal$ on $\mathbb{R}^n$ and a non-zero vector $\mathbf{y}$ in the hyperplane $\mathbb{H} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{w}^\intercal\mathbf{x} = 0\}$ determined by $\mathbf{w}$, let S be the shear operator $S = I + \mathbf{y}\mathbf{w}^\intercal$. For example, the shear operator on $\mathbb{R}^3$ for $\mathbf{w} = (1, 0, 0)$ and $\mathbf{y} = (0, 1, 0)$ is

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

  Applying the shear operator S to a vector $\mathbf{v}$ results in $\mathbf{v} + (\mathbf{w}^\intercal\mathbf{v})\mathbf{y}$, so the amount of $\mathbf{v}$ moves in direction $\mathbf{y}$ is proportional to how "far" (in a sense) $\mathbf{v}$ is from $\mathbb{H}$. So if a non-zero vector $\mathbf{v}$ is not in $\mathbb{H}$, then $S\mathbf{v}$ is not the same line as $\mathbf{v}$. An eigenvector of S must be in $\mathbb{H}$, so S can have at most $\dim(\mathbb{H}) = n - 1$ linearly independent eigenvectors. (And indeed it has exactly this many.) ∎

An $n \times n$ matrix A that has n linearly independent eigenvectors is said to be *diagonalizable*. This terminology is explained by the following proposition.

Proposition 1. *If A is an $n \times n$ diagonalizable matrix, then there exists an invertible $n \times n$ matrix $V = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ whose columns are linearly independent eigenvectors of A, and an $n \times n$ diagonal matrix Λ whose diagonal entries are corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$, such that*

$$A = V\Lambda V^{-1}.$$

Note that for the matrices V and Λ from Proposition 1, we have

$$V^{-1}AV = \Lambda,$$

a diagonal matrix. The matrix V is said to *diagonalize* A for this reason.

Proof of Proposition 1. By the Unique Linear Transformation Theorem, it suffices to show that A and $V\Lambda V^{-1}$ have the same behavior on a basis for $\mathbb{R}^n$; the basis we'll use is $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. By definition, we have $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ since $\mathbf{v}_i$ is an eigenvector of A with corresponding eigenvalue λ_i. Now we show $V\Lambda V^{-1}\mathbf{v}_i = \lambda_i\mathbf{v}_i$ as well. Since $V^{-1}V = I$, it follows that the $V^{-1}\mathbf{v}_i = \mathbf{e}_i$, the ith standard basis vector. Moreover, $\Lambda\mathbf{e}_i = \lambda_i\mathbf{e}_i$ because Λ is diagonal.[1] And finally, $V(\lambda_i\mathbf{e}_i) = \lambda_i V\mathbf{e}_i = \lambda_i\mathbf{v}_i$ since the ith column of V is $\mathbf{v}_i$. □

2 Positive semidefinite matrices

Recall that an $n \times n$ matrix A is *positive semidefinite (PSD)* if it is symmetric and $\langle \mathbf{x}, A\mathbf{x}\rangle \geq 0$ for all n-vectors $\mathbf{x}$. We also saw an equivalent definition: A is *positive semidefinite* if $A = B^\mathsf{T}B$ for some matrix B. In the course of proving the equivalence of these two definitions, we proved the following: if A is PSD, and $A = \sum_{i=1}^r \sigma_i\mathbf{u}_i\mathbf{v}_i^\mathsf{T}$ is an SVD of A, then $\mathbf{u}_i = \mathbf{v}_i$ for all $i \in \{1, \ldots, r\}$. This readily implies the following.

Theorem 1. *An $n \times n$ matrix A is PSD if and only if A has an orthonormal set of n eigenvectors for which the corresponding eigenvalues are non-negative.*

Proof. First, assume A is PSD. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ be (right) singular vectors of A from an SVD, and take $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ to be an ONB for $\mathsf{NS}(A)$. Because $\mathsf{CS}(A^\mathsf{T})$ and $\mathsf{NS}(A)$ are orthogonal complements in $\mathbb{R}^n$, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an ONB for $\mathbb{R}^n$. Using the equality of left and right singular vectors for PSD matrices, we have $A\mathbf{v}_i = \sigma_i\mathbf{v}_i$ for $i \in \{1, \ldots, r\}$, and also $A\mathbf{v}_i = \mathbf{0} = 0\mathbf{v}_i$ for $i \in \{r+1, \ldots, n\}$. This shows that the eigenvalues corresponding to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are non-negative.

Now instead assume A has an orthonormal set of n eigenvectors for which the corresponding eigenvalues are non-negative. Then by Proposition 1, we

[1] The standard coordinate basis vectors are eigenvectors of diagonal matrices.

can write $A = V\Lambda V^\top$ (since $V^{-1} = V^\top$), where V is the matrix whose columns are the n eigenvectors, and Λ is the diagonal matrix whose diagonal entries are the corresponding eigenvalues (which are non-negative). Let $B = \sqrt{\Lambda}V^\top$, so $A = B^\top B$. This shows that A is PSD. $\square$

Example. The 4×4 matrix A is equal to $B^\top B$, for A and B given below:

$$A = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}. \tag{1}$$

The matrix B has rank 3. Its singular values $\sigma_1, \sigma_2, \sigma_3$ are $2, 1, 1$, and corresponding right singular vectors are

$$\begin{aligned} \mathbf{v}_1 &= \frac{1}{2\sqrt{3}}(-3, 1, 1, 1), \\ \mathbf{v}_2 &= \frac{1}{\sqrt{12 + 6\sqrt{3}}}(0, -1-\sqrt{3}, -1, 2+\sqrt{3}), \\ \mathbf{v}_3 &= \frac{1}{\sqrt{12 + 6\sqrt{3}}}(0, -1-\sqrt{3}, 2+\sqrt{3}, -1). \end{aligned}$$

The nullspace of B is spanned by $\mathbf{v}_4 = (1/2, 1/2, 1/2, 1/2)$. It can be checked that $B^\top B\mathbf{v}_i = \sigma_i^2\mathbf{v}_i$ for each $i \in \{1, 2, 3\}$, and $B^\top B\mathbf{v}_4 = \mathbf{0}$.

Note that the choice of right singular vectors of B—and hence the eigenvectors of A—is not unique. For instance, $\mathbf{v}_2$ and $\mathbf{v}_3$ can be replaced by any two orthonormal vectors in their span. We can also replace $\mathbf{v}_1$ with $-\mathbf{v}_1$. ∎

3 Symmetric matrices

Recall that an $n \times n$ matrix A is symmetric if $A^\top = A$. It turns out that symmetric matrices behave much like PSD matrices when it comes to eigenvectors; the only difference is that the eigenvalues might be negative.

We can understand a symmetric $n \times n$ matrix A as follows:

$$A = (A - \mu I) + \mu I,$$

where μ is the smallest value of $\mathbf{u}^\mathsf{T} A\mathbf{u}$ among all unit vectors $\mathbf{u} \in \mathbb{R}^n$. The matrix $B = A - \mu I$ is PSD: if $\mathbf{x} \in \mathbb{R}^n$ is non-zero, then letting $\mathbf{u} = \mathbf{x}/\|\mathbf{x}\|$,

$$\begin{aligned} \mathbf{x}^\mathsf{T} B\mathbf{x} &= \|\mathbf{x}\|^2 \mathbf{u}^\mathsf{T}(A - \mu I)\mathbf{u} \\ &= \|\mathbf{x}\|^2(\mathbf{u}^\mathsf{T} A\mathbf{u} - \mu) \\ &\geq \|\mathbf{x}\|^2(\mu - \mu) \ = \ 0. \end{aligned}$$

The inequality follows by definition of μ.

Let V be the orthogonal matrix whose columns are the eigenvectors of B that form an ONB for $\mathbb{R}^n$, as guaranteed by Theorem 1 since B is PSD. Then V diagonalizes B; let $\Sigma = V^\mathsf{T} BV$ be the resulting diagonal matrix, whose diagonal entries are the corresponding eigenvalues, $\sigma_1, \dots, \sigma_n$. The matrix V then diagonalizes A:

$$V^\mathsf{T} AV \ = \ V^\mathsf{T}(B + \mu I)V \ = \ V^\mathsf{T} BV + \mu V^\mathsf{T} V \ = \ \Sigma + \mu I,$$

where the final step uses the fact that V is orthogonal. The matrix $\Lambda = \Sigma + \mu I$ is diagonal, with diagonal entries $\lambda_i = \sigma_i + \mu$ for $i \in \{1, \dots, n\}$. Note that while all σ_i were non-negative, it is possible that some of the $\lambda_i = \sigma_i + \mu$ are negative (as μ could be negative).

The argument above proves the following theorem.

Theorem 2. *Let A be a symmetric $n \times n$ matrix. Then A has n eigenvectors that form an ONB for $\mathbb{R}^n$.*

4 Eigenspaces

In the last example in Section 2, the fact that there are two options for $\mathbf{v}_1$ is clear: either $\mathbf{v}_1$ or $-\mathbf{v}_1$ leads to the same gain (in the context of the best fitting subspace problem). The ambiguity between which of $\mathbf{v}_1$ and $-\mathbf{v}_1$ is chosen may not seem so bad. But, perhaps more disconcerting is that in the subspace orthogonal to $\operatorname{span}(\mathbf{v}_1)$, there is an entire two-dimensional subspace in which every unit vector leads to the same gain. The difference between a one-dimensional subspace and two-dimensional subspace (of an inner product space) is critical. In a one-dimensional subspace, there are exactly two unit vectors. In a two-dimensional subspace, there are infinitely-many pairs of orthonormal vectors.

This motivates the consideration of entire subspaces of eigenvectors that correspond to the same eigenvalue. If λ is an eigenvalue of an $n \times n$ matrix A (not necessarily PSD or symmetric), then we define $E_\lambda = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \lambda\mathbf{x}\}$ to be the *eigenspace* of A corresponding to the eigenvalue λ.

Some basic properties of eigenspaces are as follows:

- The non-zero vectors of E_λ are all eigenvectors of A with λ as the corresponding eigenvalue.
- The set E_λ is a subspace of $\mathbb{R}^n$, because it is the nullspace of the matrix $A - \lambda I$. And since it contains a non-zero vector (an eigenvector), it has dimension at least one. (As we have seen, an eigenspace could have dimension more than one.)

Some extreme examples. The distinct eigenvalues of the $n \times n$ diagonal matrix A with $A_{i,i} = i$ are $1, 2, \ldots, n$; the eigenspace E_i corresponding to the eigenvalue i is $E_i = \operatorname{span}(\{\mathbf{e}_i\})$. The $n \times n$ identity matrix has only one distinct eigenvalue, 1, and its corresponding eigenspace is $E_1 = \mathbb{R}^n$. ∎

How can we understand the eigenspaces of a matrix corresponding to distinct eigenvalues? First, eigenvectors corresponding to distinct eigenvalues are linearly independent. This is intuitively true for two such eigenvectors, and an inductive argument shows that it is true for any number of eigenvectors corresponding to distinct eigenvalues.

Proposition 2. *Let A be a matrix, and suppose $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of A. Then if $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are eigenvectors of A such that λ_i is the eigenvalue corresponding to $\mathbf{v}_i$, then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent.*

A simple consequence of Proposition 2 is that the number of distinct eigenvalues of an $n \times n$ matrix is at most n, since we cannot have more than n linearly independent vectors in $\mathbb{R}^n$.

Next, if we consider a basis for each eigenspace corresponding to distinct eigenvalues, then the union of these bases is also linearly independent.

Proposition 3. *Let A be a matrix, and suppose $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of A. Let $\mathcal{B}_i$ be a basis for the eigenspace E_{λ_i} corresponding to λ_i, for each $i \in \{1, \ldots, k\}$. Then $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is linearly independent.*

Therefore, the eigenspaces $E_{\lambda_1}, \dots, E_{\lambda_k}$ corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_k$ provide a clean characterization of the behavior of $A\mathbf{x}$ for all vectors $\mathbf{x}$ in the sumset $\mathbb{W} = E_{\lambda_1} + \cdots + E_{\lambda_k}$ in the following sense: every vector $\mathbf{x} \in \mathbb{W}$ can be (uniquely) written as a sum of the "parts" of $\mathbf{x}$ from these eigenspaces, and the effect of multiplying A by $\mathbf{x}$ is simply to scale each "part" by the corresponding eigenvalue.

Theorem 3 (Eigenspace Structure Theorem)**.** *Let A be an $n \times n$ matrix. Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of A; let $E_{\lambda_i} = \{\mathbf{x} \in \mathbb{R}^n : (A - \lambda_i I)\mathbf{x} = \mathbf{0}\}$ be the eigenspace of A corresponding to its eigenvalue λ_i; and let $\mathbb{W} = E_{\lambda_1} + \cdots + E_{\lambda_k}$. For every $\mathbf{x} \in \mathbb{W}$, there are unique vectors $\mathbf{y}_1, \dots, \mathbf{y}_k$ with $\mathbf{y}_i \in E_{\lambda_i}$ for each $i \in \{1, \dots, k\}$ such that*[2]

$$\begin{aligned} \mathbf{x} &= \mathbf{y}_1 + \cdots + \mathbf{y}_k \\ \textit{and} \qquad A\mathbf{x} &= \lambda_1 \mathbf{y}_1 + \cdots + \lambda_k \mathbf{y}_k. \end{aligned}$$

(If A is diagonalizable, then $\mathbb{W} = \mathbb{R}^n$.) Furthermore, if A is symmetric, then $\mathbf{y}_i = P_i\mathbf{x}$ for each $i \in \{1, \dots, k\}$, where P_i is the orthoprojector for E_{λ_i}, so

$$\begin{aligned} I &= P_1 + \cdots + P_k \\ \textit{and} \qquad A &= \lambda_1 P_1 + \cdots + \lambda_k P_k. \end{aligned} \tag{2}$$

The decomposition of a symmetric matrix A in (2), which is unique up to the ordering of the sum, is the *spectral decomposition* of A; the distinct eigenvalues $\{\lambda_1, \dots, \lambda_k\}$ make up the *spectrum* of A.

Example. Let A be the 4×4 matrix from (1). The spectrum of A is $\{\lambda_1, \lambda_2, \lambda_3\} = \{16, 4, 0\}$, and we have

$$\begin{aligned} E_{\lambda_1} &= \operatorname{span}(\{(a, 0, 0, a) : a \in \mathbb{R}\}), \\ E_{\lambda_2} &= \operatorname{span}(\{(0, b, c, 0) : (b, c) \in \mathbb{R}^2\}), \\ E_{\lambda_3} &= \operatorname{span}(\{(a, 0, 0, -a) : a \in \mathbb{R}\}). \end{aligned}$$

Let P_1, P_2, and P_3 be the corresponding orthoprojectors for these subspaces. Then $A = 16P_{\lambda_1} + 4P_{\lambda_2} + 0P_{\lambda_3}$. ∎

[2] In terms of the concept of direct sums, we have $\mathbb{W} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

A Linear independence and eigenspaces

Proof of Proposition 2. The proof is by induction on k. In the base case $k = 1$, it is clear that $\{\mathbf{v}_1\}$ is linearly independent because, by definition, $\mathbf{v}_1 \neq \mathbf{0}$.

For some $k \geq 2$, assume as the inductive hypothesis that eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ are linearly independent, where λ_i is the eigenvalue corresponding to $\mathbf{v}_i$. Now consider an additional eigenvector $\mathbf{v}_k$ with corresponding eigenvalue λ_k, distinct from $\lambda_1, \dots, \lambda_{k-1}$. Suppose the scalar $c_1, \dots, c_k$ satisfy

$$c_1\,\mathbf{v}_1 + \cdots + c_k\,\mathbf{v}_k \;=\; \mathbf{0}. \tag{3}$$

We need to show that $c_1 = \cdots = c_k = 0$ to show that $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent.

Multiply both sides of (3) on the left by $A - \lambda_k I$. By linearity and the fact $(A - \lambda_k I)\mathbf{v}_i = (\lambda_i - \lambda_k)\mathbf{v}_i$ for each $i \in \{1, \dots, k\}$, we have

$$c_1(\lambda_1 - \lambda_k)\,\mathbf{v}_1 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\,\mathbf{v}_{k-1} \;=\; \mathbf{0}.$$

(The kth term on the left-hand side has vanished.) Since $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ are linearly independent by the inductive hypothesis, we must have

$$c_1(\lambda_1 - \lambda_k) \;=\; \cdots \;=\; c_{k-1}(\lambda_{k-1} - \lambda_k) \;=\; 0.$$

The distinctness of $\lambda_1, \dots, \lambda_k$ implies that $c_i = 0$ for each $i \in \{1, \dots, k-1\}$. Therefore, (3) becomes $c_k\mathbf{v}_k = \mathbf{0}$. Since $\mathbf{v}_k$ is non-zero, this equation can only hold if $c_k = 0$. So we have shown that $c_1 = \cdots = c_k = 0$. This implies that $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent, which proves the inductive step. So the overall claim follows by the principle of mathematical induction. □

A corollary of Proposition 2 is the following.

Corollary 1. *Let A be a matrix, and suppose $\lambda_1, \dots, \lambda_k$ are distinct eigenvalues of A. Then if vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ satisfy $\mathbf{v}_i \in E_{\lambda_i}$ for each $i \in \{1, \dots, k\}$ and $\mathbf{v}_1 + \cdots + \mathbf{v}_k = \mathbf{0}$, then $\mathbf{v}_1 = \cdots = \mathbf{v}_k = \mathbf{0}$.*

Proof of Proposition 3. Let $\mathcal{B}_i = \{\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,n_i}\}$, where $n_i = |\mathcal{B}_i|$. Suppose there are scalars $c_{i,j}$ for $i \in \{1, \dots, k\}$ and $j \in \{1, \dots, n_i\}$ such that

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} c_{i,j}\mathbf{v}_{i,j} \;=\; \mathbf{0}.$$

We need to show that all of these $c_{i,j}$'s must be zero.

Define $\mathbf{w}_i = \sum_{j=1}^{n_i} c_{i,j}\mathbf{v}_{i,j}$, so $\mathbf{w}_i \in E_{\lambda_i}$. Then we have $\mathbf{w}_1 + \cdots + \mathbf{w}_k = \mathbf{0}$. By Corollary 1, we must have $\mathbf{w}_1 = \cdots = \mathbf{w}_k = \mathbf{0}$, i.e., for each i,

$$\sum_{j=1}^{n_i} c_{i,j}\mathbf{v}_{i,j} \;=\; \mathbf{0}.$$

But $\mathbf{v}_{i,1}, \ldots, \mathbf{v}_{i,n_i}$ are linearly independent since $\mathcal{B}_i$ is a basis, so $c_{i,1} = \cdots = c_{i,n_i} = 0$. This holds for all i, and thus all of these scalars $c_{i,j}$ must be zero. So we have proved the linear independence of $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$. □

Proof of Theorem 3. We just prove the first part of the claim, since the "Furthermore" part for symmetric matrices is immediate after combining with Theorem 2.

For each $i \in \{1, \ldots, k\}$, let $\mathcal{B}_i = (\mathbf{v}_{i,1}, \ldots, \mathbf{v}_{i,n_i})$ be an ordered basis for E_{λ_i}, where $n_i = \dim E_{\lambda_i}$. By Proposition 3, $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is an ordered basis for $\mathbb{W} = E_{\lambda_1} + \cdots + E_{\lambda_k}$. Therefore, for every $\mathbf{x} \in \mathbb{W}$, there is a unique choice of scalars $c_{1,1}, \ldots, c_{1,n_1}, \ldots, c_{k,1}, \ldots, c_{k,n_k}$ such that

$$\mathbf{x} \;=\; \mathbf{y}_1 + \cdots + \mathbf{y}_k$$

where, for each $i \in \{1, \ldots, k\}$,

$$\mathbf{y}_i \;=\; c_{i,1}\,\mathbf{v}_{i,1} + \cdots + c_{i,n_i}\,\mathbf{v}_{i,n_i}.$$

Since $\mathbf{y}_i \in E_{\lambda_i}$, it follows that $A\mathbf{y}_i = \lambda_i\mathbf{y}_i$. And therefore, by linearity,

$$A\mathbf{x} \;=\; \lambda_1\,\mathbf{y}_1 + \cdots + \lambda_k\,\mathbf{y}_k.$$

It remains to show that the choice of $\mathbf{y}_1, \ldots, \mathbf{y}_k$ is unique. First, we prove that for every $i \in \{1, \ldots, k\}$,

$$E_{\lambda_i} \cap \left(\sum_{j \neq i} E_{\lambda_j} \right) \;=\; \{\mathbf{0}\}.$$

Suppose $\mathbf{v}$ is a vector in the intersection. Write $\mathbf{v}$ as a linear combination of basis vectors from $\mathcal{B}_i$,

$$\mathbf{v} \;=\; c_{i,1}\,\mathbf{v}_{i,1} + \cdots + c_{i,n_i}\,\mathbf{v}_{i,n_i},$$

and then write $\mathbf{v}$ as a linear combination of basis vectors from $\bigcup_{j \neq i} \mathcal{B}_j$,

$$\mathbf{v} = \sum_{j \neq i} c_{j,1} \mathbf{v}_{j,1} + \cdots + c_{j,n_j} \mathbf{v}_{j,n_j}.$$

The difference is $\mathbf{0}$, and now $\mathbf{0}$ expressed as a linear combination of basis vectors from $\mathcal{B}$:

$$\mathbf{0} = c_{i,1} \mathbf{v}_{i,1} + \cdots + c_{i,n_i} \mathbf{v}_{i,n_i} - \sum_{j \neq i} c_{j,1} \mathbf{v}_{j,1} + \cdots + c_{j,n_j} \mathbf{v}_{j,n_j}.$$

Since $\mathcal{B}$ is linearly independent, all of the coefficients in the linear combination must be zero. This we conclude that $\mathbf{v} = \mathbf{0}$.

Now suppose $\mathbf{x} = \mathbf{z}_1 + \cdots + \mathbf{z}_k$ where $\mathbf{z}_i \in E_{\lambda_i}$ for each $i \in \{1, \dots, k\}$. We show that $\mathbf{z}_i = \mathbf{y}_i$ for each $i \in \{1, \dots, k\}$. Since $\mathbf{z}_1 + \cdots + \mathbf{z}_k = \mathbf{y}_1 + \cdots + \mathbf{y}_k$, it follows that for every $i \in \{1, \dots, k\}$, we have

$$\mathbf{z}_i - \mathbf{y}_i = \sum_{j \neq i} \mathbf{y}_j - \mathbf{z}_j.$$

The left-hand side is in E_{λ_i}, while the right-hand side is in $\sum_{j \neq i} E_{\lambda_j}$. But the intersection of these two subspaces is $\{\mathbf{0}\}$, as argued above. Hence we conclude that $\mathbf{z}_i = \mathbf{y}_i$ for every $i \in \{1, \dots, k\}$.

Finally, it is clear that if A is diagonalizable, then $|\mathcal{B}| = n$ by Proposition 3, and hence $\mathbb{W} = \mathbb{R}^n$ by the Basis Sufficiency Theorem. $\square$

B The determinant and characteristic polynomial of a matrix

There is a function $\det \colon \mathsf{M}_{n \times n}(\mathbb{R}) \to \mathbb{R}$ called the *determinant* that is commonly used to derive a test of diagonalizability. (It is also used for other things, such as computing volumes of parallelepipeds, which is needed to perform a change-of-variables in multiple integration.)

Here are some key properties of the determinant function:

1. If B is obtained from A by swapping two rows, then $\det(B) = -\det(A)$.

2. If B is obtained from A by multiplying a row by the scalar c, then $\det(B) = c \det(A)$.
3. If B is obtained from A by adding a multiple of row i to row j, where $i \neq j$, then $\det(B) = \det(A)$.
4. $\det(I) = 1$.
5. $\det(AB) = \det(A)\det(B)$.
6. A is invertible if and only if $\det(A) \neq 0$.

(The last two properties can be derived from the first four.)

Using the definition of eigenvalue, the Invertibility Theorem, and the final property of determinants given above, we find that the following statements are equivalent:

- λ is an eigenvalue of A;
- $A\mathbf{v} = \lambda\mathbf{v}$ for some non-zero vector $\mathbf{v}$;
- $(A - \lambda I)\mathbf{v} = \mathbf{0}$ for some non-zero vector $\mathbf{v}$;
- $A - \lambda I$ is not invertible;
- $\det(A - \lambda I) = 0$.

For an $n \times n$ matrix A, if t is treated as a variable, then $\det(A - tI)$ is a polynomial of degree n in the variable t, with leading coefficient $(-1)^n$. This polynomial is called the *characteristic polynomial* of A. The roots of the characteristic polynomial of A are the eigenvalues of A. This polynomial is said to *split* if it can be factored as

$$\det(A - tI) \;=\; (-1)^n (t - \lambda_1)^{d_1} \cdots (t - \lambda_k)^{d_k},$$

where $\lambda_1, \dots, \lambda_k$ are distinct scalars, and $d_1, \dots, d_k$ are positive integers that sum to n. The distinct scalars $\lambda_1, \dots, \lambda_k$ are the distinct roots of the characteristic polynomial, and hence they are the distinct eigenvalues of A. The number d_i is the *algebraic multiplicity* of λ_i. The dimension of the eigenspace E_{λ_i} is called the *geometric multiplicity* of λ_i; it is an integer between 1 and the algebraic multiplicity of λ_i. In order to have n linearly independent eigenvectors, it must be the case that the geometric multiplicity is equal to the algebraic multiplicity, for each of the distinct eigenvalues.

Theorem 4. *A square matrix is diagonalizable if and only if (i) its characteristic polynomial splits, and (ii) for each of the distinct eigenvalues of the matrix, the geometric multiplicity is equal to the algebraic multiplicity.*

Example. Let A be the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

The characteristic polynomial of A is

$$\det\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} - t\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1-t & 0 \\ 1 & 1-t \end{bmatrix}\right) = (t-1)^2.$$

This polynomial splits. It has a single root (eigenvalue), 1, with algebraic multiplicity 2. The eigenspace of A corresponding to this eigenvalue is

$$\begin{aligned} E_1 &= \{\mathbf{x} \in \mathbb{R}^2 : (A - I)\mathbf{x} = \mathbf{0}\} \\ &= \left\{(x_1, x_2) \in \mathbb{R}^2 : \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\} \\ &= \{(0, x_2) : x_2 \in \mathbb{R}\}. \end{aligned}$$

The dimension of E_1 is 1, so the geometric multiplicity of the eigenvalue 1 is $1 \neq 2$. Hence A is not diagonalizable. ∎