Lecture 8: Eigenvalues, eigenvectors and diagonalization

Francesco Preta

July 30, 2020

In this lecture we will develop the theory of eigendecomposition for square matrices which has many useful applications, including dynamical systems and Markov chains.

Eigenvalues and Eigenvectors

Definition 1. Let A be a $n \times n$ matrix. We say that λ is an eigenvalue of A with eigenvector $\mathbf{v} \neq \mathbf{0}$ if

$$A\mathbf{v} = \lambda \mathbf{v}$$

Eigenvectors of a matrix A are vectors in $\mathbb{R}^n$ on which A acts by scalar multiplication.

Example: consider $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Then we have

$$\begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

In this case, $\mathbf{v}$ is an eigenvector for A with eigenvalue 5.

In general, an eigenvalue of a matrix can have more than one eigenvector.

Definition 2. Let A be a matrix and λ be one of its eigenvalues. The space spanned by all the eigenvectors for λ is called the *eigenspace* of λ, denoted by $E(\lambda)$. Its dimension (corresponding to the number of linearly independent eigenvectors for λ) is called the *geometric multiplicity* of λ.

Example: consider the matrix

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

and the vectors $\mathbf{v}_1 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$. We have that

$$A\mathbf{v}_1 = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} = -\mathbf{v}_1$$
$$A\mathbf{v}_1 = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = -\mathbf{v}_2$$

so that $\lambda = -1$ is an eigenvalue of A with geometric multiplicity at least 2 and the eigenspace of $\lambda = -1$ is such that

$$\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2) = \left\{ \begin{bmatrix} \alpha \\ -2\alpha - 2\beta \\ \beta \end{bmatrix} \mid \alpha, \beta \in \mathbb{R} \right\} \subset E(-1).$$

We haven't yet found a way to prove that these two eigenvectors are actually a basis for $E(-1)$, so that the inclusion above is actually an equality, but we will now develop a method to show that it is indeed the case.

Notice that $\lambda = 0$ can be an eigenvalue of a matrix A, in which case the associated eigenspace is merely the Kernel of A.

The characteristic equation

We have defined eigenvalues and eigenvectors but we need to develop an algorithm to find them. Notice that for a given matrix A, we need to find λ and $\mathbf{v} \neq \mathbf{0}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. This is equivalent to solving the equation

$$(A - \lambda I)\mathbf{v} = \mathbf{0}$$

which corresponds to finding the elements of $\operatorname{Ker}(A - \lambda I)$. However, remember that λ is also an unknown in this equation, so that we first need to find the λ for which $A - \lambda I$ has non-trivial kernel. This happens when

$$\det(A - \lambda I) = 0 \tag{1}$$

Equation (1) is called the *characteristic equation* for a matrix A. Its solutions determine the eigenvalues of A.

Example: consider the matrix of our previous example:

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Then

$$A - \lambda I = \begin{bmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{bmatrix}$$

and

$$\det(A - \lambda I) = -\lambda(3-\lambda)^2 + 16 + 16 + 16\lambda + 8(\lambda - 3) = -\lambda^3 + 6\lambda^2 + 15\lambda + 8$$

so the characteristic equation is

$$\lambda^3 - 6\lambda^2 - 15\lambda - 8 = 0$$

We know that $\lambda = -1$ is a solution to this equation, so that by dividing for $\lambda + 1$ we get

$$(\lambda+1)(\lambda^2 - 7\lambda - 8) = 0$$

which can be factored as

$$(\lambda+1)(\lambda+1)(\lambda-8) = (\lambda+1)^2(\lambda-8)$$

so that the eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = 8$.

In general, the characteristic equation for a $n \times n$ matrix A is a polynomial of degree n. That means that it admits n (possibly complex) solutions, considering possible multiplicity. That is, we can factor it in the following way:

$$\det(A - \lambda I) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{m_i}$$

so that $k \leq n, m_i \geq 1$ and $\sum_{i=1}^{k} m_i = n$. The value m_i is called the *algebraic multiplicity* of a solution λ_i. In the previous example, -1 has algebraic multiplicity 2, while 8 has algebraic multiplicity 1.

Before we delve into the specifics of the relation between algebraic and geometric multiplicity, let us develop a method to find the eigenvectors corresponding to given eigenvalues.

In order to do so, for each eigenvalue λ_i for $i = 1, ..., k$ that we have found, consider the homogeneous system

$$(A - \lambda_i I)\mathbf{v} = \mathbf{0} \tag{2}$$

and solve it for $\mathbf{v}$. Since $\det(A - \lambda_i I) = 0$, the system will have infinitely many solutions (and not only the trivial one, $\mathbf{v} = \mathbf{0}$). Therefore the set of solutions will be a vector space spanned by a minimal set of eigenvectors. The dimension of this space will be found by looking at the solution set of Equation (2).

In the previous example, for $\lambda_1 = -1$ we need to solve the system

$$\begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which can be rewritten as

$$\begin{cases} 4v_1 + 2v_2 + 4v_3 = 0 \\ 2v_1 + v_2 + 2v_3 = 0 \\ 4v_1 + 2v_2 + 4v_3 = 0 \end{cases}$$

which corresponds to the solution set given by only one of the three equations (the others are multiples). By choosing v_1 and v_3 as free variables, we get

$$E(-1) = \left\{ \begin{bmatrix} \alpha \\ -2\alpha - 2\beta \\ \beta \end{bmatrix} \mid \alpha, \beta \in \mathbb{R} \right\}$$

and notice that for $\alpha = 1, \beta = 0$ and $\alpha = 0, \beta = 1$ we obtain the same eigenvectors as before.

As for $\lambda_2 = 8$, we have the system

$$\begin{bmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which can be rewritten as

$$\begin{cases} -5v_1 + 2v_2 + 4v_3 = 0 \\ 2v_1 - 8v_2 + 2v_3 = 0 \\ 4v_1 + 2v_2 - 5v_3 = 0 \end{cases}$$

We can use Gaussian elimination to get an echelon form of the type

$$\begin{bmatrix} 1 & -4 & 1 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and a reduced echelon form of the type

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so that v_3 is the free variable and we get the eigenspace to be

$$E(8) = \left\{ \begin{bmatrix} \alpha \\ \frac{\alpha}{2} \\ \alpha \end{bmatrix} \mid \alpha \in \mathbb{R} \right\}$$

For instance by choosing $\alpha = 2$ we get the eigenvector $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$.

Notice that we have made extensive use of the following Theorem:

Theorem 1. *Let λ be an eigenvalue for a square matrix $A \in \mathbb{R}^{n \times n}$. Then the eigenspace $E(\lambda)$ is a vector subspace of $\mathbb{R}^n$.*

Proof. By definition, $E(\lambda) = \operatorname{Ker}(A - \lambda I)$ which is always a vector subspace of $\mathbb{R}^n$. □

Moreover, notice that the $E(0) = \operatorname{Ker}(A - 0I) = \operatorname{Ker}(A)$ which is non-trivial if and only if $\det(A) = 0$.

Multiplicity and diagonalization

In the previous sections we have defined two important concepts: the algebraic multiplicity, which represents the number of times a given eigenvalues solves the characteristic equation, and the geometric multiplicity, which represents the dimension of the corresponding eigenspace. In general, the geometric multiplicity is less than or equal to the algebraic multiplicity. When they coincide on every eigenvector, we can find a very useful decomposition of a square matrix: its diagonalization.

Example: consider the matrix

$$A = \begin{bmatrix} 5 & 2 & 4 \\ 0 & 4 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

In this case the characteristic equation corresponds to

$$(5 - \lambda)(4 - \lambda)(5 - \lambda) = 0$$

so that 5 is an eigenvalue of A with algebraic multiplicity 2. Its eigenvectors solve the equations

$$\begin{cases} 2v_2 + 4v_3 = 0 \\ -v_2 + v_3 = 0 \\ v_1 \text{ free} \end{cases}$$

The only solution to this equation is given by multiples of $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, so that

$$E(5) = \{\alpha \mathbf{e}_1 \,|\, \alpha \in \mathbb{R}\}$$

The degrees of freedom of this set of solutions is only one, so the dimension of $E(5)$ is 1. This means that the geometric multiplicity is strictly less than the algebraic multiplicity.

The reason why we would like algebraic and geometric multiplicity to coincide is because if they do we can find n linearly independent eigenvectors that span the entire space $\mathbb{R}^n$. This happens because of the following:

Theorem 2. *Let $\mathbf{v}_1 \in E(\lambda_1)$ and $\mathbf{v}_2 \in E(\lambda_2)$ be two nontrivial eigenvectors for a square matrix $A \in \mathbb{R}^{n \times n}$, with λ_1 and λ_2 two different eigenvalues of A. Then $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.*

Proof. By contradiction, suppose there exists α such that $\mathbf{v}_1 = \alpha\mathbf{v}_2$. Then

$$\lambda_1\mathbf{v}_1 = A\mathbf{v}_1 = A\alpha\mathbf{v}_2 = \alpha\lambda_2\mathbf{v}_2 = \lambda_2\mathbf{v}_1$$

But then $\lambda_1 = \lambda_2$ which contradicts our assumption. $\square$

If there are n eigenvectors of A, $\mathbf{v}_1, ..., \mathbf{v}_n$, that are linearly independent, then we can apply a change of basis, from the canonical basis to $\mathbf{v}_1, ..., \mathbf{v}_n$ to observe the action of A on these eigenspaces. In particular, for any $\mathbf{x} = \sum_{i=1}^n \alpha_i\mathbf{v}_i$, for $\mathbf{v}_i \in E(\lambda_i)$, we get

$$A\mathbf{x} = \sum_{i=1}^n \alpha_i A\mathbf{v}_i = \sum_{i=1}^n \alpha_i\lambda_i\mathbf{v}_i$$

This shows that in a basis made of eigenvectors, A acts by multiplying each one of them by a constant. This is the same type of action of a diagonal matrix on the canonical basis, therefore it induces a *diagonalization* of A.

Theorem 3. *Let A be a square $n \times n$ matrix with n independent eigenvectors. Then there exists an invertible matrix P and a diagonal matrix Λ such that $A = P\Lambda P^{-1}$. For a given ordering of n independent eigenvectors $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$, the matrix P is a matrix having i-th column equal to $\mathbf{v}_i$ and the matrix Λ has on the i-th element of the diagonal the eigenvalue λ_i corresponding to $\mathbf{v}_i$.*

Example: consider the matrix of the previous example

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

which we found to have eigenvalues $\lambda_1 = -1$ with multiplicity 2 and corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$ and $\lambda_2 = 8$ with eigenvectors $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$. Then we have

$$P = \begin{bmatrix} 1 & 0 & 2 \\ -2 & -2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \qquad \Lambda = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Then we can calculate P^{-1} through the inverse matrix algorithm by finding the reduced echelon form of the matrix:

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ -2 & -2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{bmatrix}$$

This corresponds to

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{9} & -\frac{2}{9} & -\frac{4}{9} \\ 0 & 1 & 0 & -\frac{4}{9} & -\frac{2}{9} & \frac{5}{9} \\ 0 & 0 & 1 & \frac{2}{9} & \frac{1}{9} & \frac{2}{9} \end{array}\right]$$

so that

$$P^{-1} = \begin{bmatrix} \frac{5}{9} & -\frac{2}{9} & -\frac{4}{9} \\ -\frac{4}{9} & -\frac{2}{9} & \frac{5}{9} \\ \frac{2}{9} & \frac{1}{9} & \frac{2}{9} \end{bmatrix}$$

So now consider

$$\begin{bmatrix} 1 & 0 & 2 \\ -2 & -2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{bmatrix} \begin{bmatrix} \frac{5}{9} & -\frac{2}{9} & -\frac{4}{9} \\ -\frac{4}{9} & -\frac{2}{9} & \frac{5}{9} \\ \frac{2}{9} & \frac{1}{9} & \frac{2}{9} \end{bmatrix} = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

The three matrices on the left corresponds to the diagonalization of A.

The meaning of this decomposition can be seen as such: P^{-1} is the change of basis matrix from the canonical basis $\{\mathbf{e}_i\}_{i=1}^n$ to the basis of eigenvectors $\{\mathbf{v}_i\}_{i=1}^n$. That means that if $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i = \sum_{i=1}^n \alpha_i\mathbf{v}_i$, then

$$P\Lambda P^{-1}\sum_{i=1}^n \alpha_i\mathbf{v}_i = \sum_{i=1}^n \alpha_i P\Lambda\mathbf{e}_i = \sum_{i=1}^n \alpha_i\lambda_i P\mathbf{e}_i = \sum_{i=1}^n \alpha_i\lambda_i\mathbf{v}_i = A\sum_{i=1}^n \alpha_i\mathbf{v}_i$$

so that A and $P\Lambda P^{-1}$ coincide for every $\mathbf{x} \in \mathbb{R}^n$.

How do we know if a matrix A can be diagonalized? We need conditions so that we can find a basis of $\mathbb{R}^n$ made of n independent eigenvectors. We have the following theorem:

Theorem 4. *Let A be a $n\times n$ matrix. Then A is diagonalizable if and only if for every eigenvalue λ_i of A, the algebraic multiplicity and the geometric multiplicity coincide. In particular, if A has all distinct eigenvalues, it is diagonalizable.*

Example: consider the matrix

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

is it diagonalizable? If yes, find the diagonalization. If not, explain why.

Step 1: find the eigenvalues. Since A is a triangular matrix, the characteristic equation is given by

$$-\lambda(2-\lambda)(1-\lambda) = 0$$

so that the eigenvalues are $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_3 = 2$. Since all the eigenvalues are different, the matrix is diagonalizable.

Step 2: find the corresponding eigenvectors.

$\lambda_1 = 0$ gives us the system

$$\begin{cases} -2v_2 + v_3 = 0 \\ 2v_2 + v_3 = 0 \\ v_3 = 0 \end{cases}$$

which tells us that

$$E(0) = \{\alpha \mathbf{e}_1 \,|\, \alpha \in \mathbb{R}\}$$

$\lambda_2 = 1$ gives us the system:

$$\begin{cases} -v_1 - 2v_2 + v_3 = 0 \\ v_2 + v_3 = 0 \end{cases}$$

which tells us that

$$E(1) = \left\{ \begin{bmatrix} 3\beta \\ -\beta \\ \beta \end{bmatrix} \,|\, \beta \in \mathbb{R}\right\}$$

Finally, $\lambda_3 = 2$ gives us the system

$$\begin{cases} -2v_1 - 2v_2 + v_3 = 0 \\ v_3 = 0 \\ v_3 = 0 \end{cases}$$

which tells us that

$$E(2) = \left\{ \begin{bmatrix} \gamma \\ -\gamma \\ 0 \end{bmatrix} \,|\, \gamma \in \mathbb{R}\right\}$$

At this point, we would always be able to find out if the matrix is diagonalizable, depending on the multiplicities of its eigenvalues. If it is diagonalizable, proceed to Step 3.

Step 3: write down the matrices that compose the diagonalization. First we choose n independent eigenvectors:

$$\mathbf{v}_1 = \mathbf{e}_1 \qquad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

so that we have

$$P = \begin{bmatrix} 1 & 3 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \Lambda = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Then find P^{-1}. If we use the inverse matrix algorithm on P we need to find the reduced echelon form of the following matrix:

$$\begin{bmatrix} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

which is given by

$$\begin{bmatrix} 1 & 0 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & -1 \end{bmatrix}$$

so that

$$P^{-1} = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}$$

and we can write

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & -2 \\ 0 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}$$

A useful application of diagonalization is the following: if we need to calculate a power of a given matrix A, say A^k, we can avoid computing $k-1$ matrix multiplications, but rather we can diagonalize it to obtain:

$$A^k = (P\Lambda P^{-1})^k = P\Lambda P^{-1} \cdot P\Lambda P^{-1} \cdot ... \cdot P\Lambda P^{-1} = P\Lambda^k P^{-1}$$

and notice that computing Λ^k corresponds only to raise each element on the diagonal to the k-th power. This also tells us that for every $k \in \mathbb{Z}$, A^k is diagonalizable with eigenvalues $\lambda_1^k, ..., \lambda_n^k$. Notice that if at least one of them is zero, then A is not invertible (it has nontrivial kernel) and therefore it doesn't make sense to talk about negative powers of A.

Complex eigenvalues

Sometimes the characteristic equation of a matrix having real coefficients doesn't have solutions in $\mathbb{R}$. For example, consider the matrix

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

The characteristic equation $\det(A - \lambda I) = 0$ corresponds to the polynomial

$$\lambda^2 + 1 = 0$$

which doesn't have solutions in $\mathbb{R}$. However, it has solutions in $\mathbb{C}$, as $\lambda_1 = i$ and $\lambda_2 = -i$. In general, the following is always true:

Theorem 5 (Fundamental Theorem of Algebra)**.** *Let p be a polynomial of degree n. Then $p(\lambda) = 0$ admits n solutions in $\mathbb{C}$, possibly with multiplicity. In particular, if p has real coefficients and $\lambda_0 \in \mathbb{C} \setminus \mathbb{R}$ is a solution to $p(\lambda) = 0$, then its complex conjugate $\bar{\lambda}_0$ is also a solution to $p(\lambda) = 0$.*

In particular, the characteristic polynomial of a $n \times n$ matrix with real coefficients is always a polynomial of degree n. It will admit n solutions, up to multiplicity and could potentially admit pairs of conjugate complex solutions. In this case, diagonalization is not really effective, since the eigenvectors will also be in $\mathbb{C}$. In our previous example, the eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

In the 2×2 case we can actually consider another type of decomposition that gives us more information: for any A having a pair of complex eigenvalues, we want to find a conjugate of the form

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

in this way, for $r = \sqrt{a^2 + b^2}$, we can decompose it into

$$C = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \frac{a}{r} & -\frac{b}{r} \\ \frac{b}{r} & \frac{a}{r} \end{bmatrix}$$

and treat the first matrix as a dilatation by r along both axis and the second as a rotation by a θ such that $\cos(\theta) = \frac{a}{r}$ and $\sin(\theta) = \frac{b}{r}$. Then for a given $A \in \mathbb{R}^{2\times 2}$ having complex eigenvalues, let $\mathbf{v}$ be one of the two eigenvectors and take

$$S = \begin{bmatrix} \mathrm{Re}(\mathbf{v}) & \mathrm{I}(\mathbf{v}) \end{bmatrix}$$

where $\mathrm{Re}(\mathbf{v})$ is the real part and $\mathrm{I}(\mathbf{v})$ is the imaginary part of $\mathbf{v}$. Then there exists a C such that

$$A = SCS^{-1}$$

Example:

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{3}{5} \\ \frac{3}{4} & \frac{11}{10} \end{bmatrix}$$

Then the characteristic equation is $5\lambda^2 - 8\lambda + 5 = 0$ and the eigenvalues are $\lambda_{1,2} = \frac{4}{5} \pm \frac{3}{5}i$. Then the eigenvector for λ_1 is

$$\mathbf{v}_1 = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$$

so we can consider

$$S = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}, \qquad S^{-1} = \frac{1}{20} \begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}$$

so that

$$C = S^{-1}AS = \frac{1}{20} \begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{3}{5} \\ \frac{3}{4} & \frac{11}{10} \end{bmatrix} \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix}$$

and we can see C as a pure rotation matrix since $\det(C) = 1$.

The decomposition $A = SCS^{-1}$ works as follows: S represents a change of basis matrix, C represents a rotation matrix (possibly with dilatation by the same amount in both directions) and S^{-1} represents a return to the original basis. To this extent, every 2×2 matrix with complex eigenvalues is, in some sense, a rotation.

Spectral properties

We list here a few properties that can be of use when dealing with eigenvalues and eigenvectors.

Definition 3. Let A be a $n \times n$ matrix, then we define the *Trace* of A to be

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{i,i}$$

An interesting property of trace and determinant is that they are invariant under conjugation:

Theorem 6. *Let A be a $n \times n$ matrix and Λ the diagonal matrix with its eigenvalues, then*

- $\mathrm{Tr}(A) = \mathrm{Tr}(\Lambda) = \sum_{i=1}^{n} \lambda_i$.
- $\det(A) = \det(\Lambda) = \prod_{i=1}^{n} \lambda_i$.
- $\mathrm{Rank}(A) = \mathrm{Rank}(\Lambda) =$ *number of nonzero eigenvalues of A.*

The theorem can be used both ways: if we already have an eigendecomposition, we can calculate Trace, Rank and Determinant of a matrix through its eigenvalues. On the other hand, we can calculate eigenvalues of 2×2 matrix A just by considering $\mathrm{Tr}(A)$ and $\det(A)$.

Example: consider the matrix

$$A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}$$

We have that $\mathrm{Tr}(A) = 0$ and $\det(A) = -4$. Therefore we need to find two eigenvalues whose product is -3 and whose sum is 0. This is satisfied by $\lambda_1 = 2$ and $\lambda_2 = -2$.

Another interesting property is the following:

Theorem 7. *Let A be a symmetric $n \times n$ matrix. Then A admits a diagonalization by n orthonormal eigenvectors.*

In order to see that if A is symmetric different eigenspaces have orthogonal eigenvalues, let $\mathbf{v}_1 \in E(\lambda_1)$ and $\mathbf{v}_2 \in E(\lambda_2)$ for $\lambda_1 \neq \lambda_2$. Then

$$\lambda_1\langle \mathbf{v}_1, \mathbf{v}_2\rangle = \langle A\mathbf{v}_1, \mathbf{v}_2\rangle = \langle \mathbf{v}_1, A\mathbf{v}_2\rangle = \lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle$$

which is possible if and only if $\langle \mathbf{v}_1, \mathbf{v}_2\rangle = 0$.

Example: consider

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

then the characteristic polynomial is given by

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & -2 & 2 \\ -2 & -2-\lambda & 4 \\ 2 & 4 & -2-\lambda \end{vmatrix} = -\lambda^3 - 3\lambda^2 + 24\lambda - 28$$

which can be factored as

$$p(\lambda) = -(\lambda - 2)^2(\lambda + 7)$$

Now we can find eigenvectors for $E(2)$ by solving the following system:

$$\begin{bmatrix} -1 & -2 & 2 & 0 \\ -2 & -4 & 4 & 0 \\ 2 & 4 & -4 & 0 \end{bmatrix}$$

Notice that the second and the third rows are multiples of the first, so the solutions have form

$$\mathrm{Ker}(A - 2I) = \Big\{ \begin{bmatrix} -2\alpha + 2\beta \\ \alpha \\ \beta \end{bmatrix} \mid \alpha, \beta \in \mathbb{R} \Big\}$$

so that we can pick a basis $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ and apply Gram-Schmidt to find

$$\mathbf{u}_1 = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}$$

and then

$$\mathbf{u}_2' = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \langle \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix} \rangle \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \begin{bmatrix} \frac{1}{2\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{\sqrt{5}}{4} \end{bmatrix}$$

Finally, we find $E(-7)$ by considering the solutions to the system having augmented matrix

$$\begin{bmatrix} 8 & -2 & 2 & 0 \\ -2 & 5 & 4 & 0 \\ 2 & 4 & 5 & 0 \end{bmatrix}$$

which can be solved to get an echelon form

$$\begin{bmatrix} 8 & -2 & 2 & 0 \\ 0 & 18 & 18 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and a reduced echelon form

$$\begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so that

$$Ker(A+7I) = \left\{ \begin{bmatrix} -\frac{\gamma}{2} \\ -\gamma \\ \gamma \end{bmatrix} \mid \gamma \in \mathbb{R} \right\}$$

and we can get the element of unit norm in this one-dimensional vector space to be

$$\mathbf{u}_3 = \begin{bmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

and notice that $\langle \mathbf{u}_1, \mathbf{u}_3 \rangle = 0$ and $\langle \mathbf{u}_2, \mathbf{u}_3 \rangle = 0$.

Finally, notice that the diagonalization is now easier to compute since if U is the matrix having the i-th unitary eigenvector as i-th column, then $U^{-1} = U^T$, so that $A = U \Lambda U^T$.