Eigenvalues and eigenvectors

Definition: For a vector space V over a field $\mathbb{K}$ and a linear map $f : V \to V$, the *eigenvalue* of f is any $\lambda \in \mathbb{K}$ for which exists a vector $\boldsymbol{u} \in V \setminus \mathbf{0}$ such that $f(\boldsymbol{u}) = \lambda \boldsymbol{u}$.

The *eigenvector* corresponding to an eigenvalue λ is any vector $\boldsymbol{u}$ such that $f(\boldsymbol{u}) = \lambda \boldsymbol{u}$.

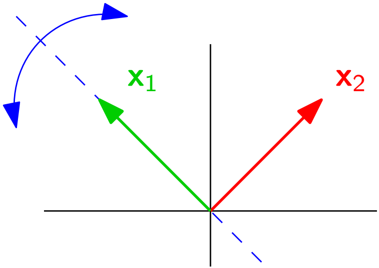
If V is of finite dimension n, then f can be represented by the matrix $\boldsymbol{A} = [f]_{XX} \in \mathbb{K}^{n \times n}$ w.r.t. some basis X of V.

This way we get *eigenvalues* $\lambda \in \mathbb{K}$ *and eigenvectors* $\boldsymbol{x} \in \mathbb{K}^n$ *of matrices* — these shall satisfy $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$.

The collection of all eigenvalues of a matrix is its *spectrum*.

Examples — a linear map in the plane $\mathbb{R}^2$

The axis symmetry by the axis of the 2nd and 4th quadrant

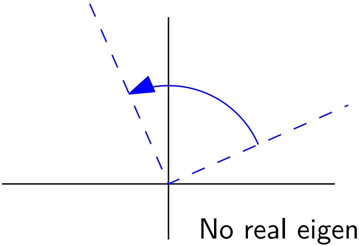


$$[f]_{KK} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

$\lambda_1 = 1$ $\qquad$ $\boldsymbol{x}_1 = c \cdot (-1, 1)^T$

$\lambda_2 = -1$ $\qquad$ $\boldsymbol{x}_2 = c \cdot (1, 1)^T$

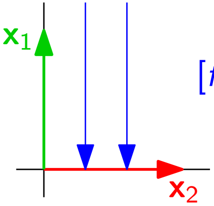
The rotation by the right angle



$$[f]_{KK} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

No real eigenvalues nor eigenvectors exist.

The projection onto the first coordinate

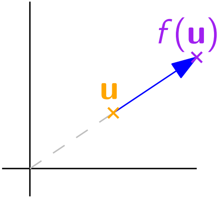


$$[f]_{KK} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\lambda_1 = 0 \qquad \boldsymbol{x}_1 = c \cdot (0, 1)^T$$

$$\lambda_2 = 1 \qquad \boldsymbol{x}_2 = c \cdot (1, 0)^T$$

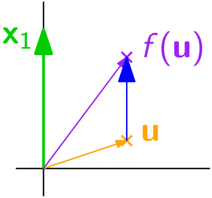
Scaling by the factor 2



$$[f]_{KK} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$\lambda_1 = 2$ *Every vector* is an eigenvector.

A linear map given by a matrix



$$[f]_{KK} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$\lambda_1 = 1$ $\boldsymbol{x}_1 = c \cdot (0, 1)^T$

Eigenvectors and eigenvalues of a diagonal matrix $\boldsymbol{D}$

The equation

$$\boldsymbol{D}\boldsymbol{x} = \begin{pmatrix} d_{1,1} & 0 & \dots & 0 \\ 0 & d_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} d_{1,1}x_1 \\ d_{2,2}x_2 \\ \vdots \\ d_{n,n}x_n \end{pmatrix} = \lambda \boldsymbol{x}$$

is solved by the following eigenvalues and eigenvectors:

$$\begin{gathered} \lambda = d_{1,1} \text{ and } \boldsymbol{x} = \boldsymbol{e}^1 = (1, 0, 0, \dots, 0)^T, \\ \lambda = d_{2,2} \text{ and } \boldsymbol{x} = \boldsymbol{e}^2 = (0, 1, 0, \dots, 0)^T, \\ \vdots \\ \lambda = d_{n,n} \text{ and } \boldsymbol{x} = \boldsymbol{e}^n = (0, 0, \dots, 0, 1)^T. \end{gathered}$$

Hence the eigenvalues of $\boldsymbol{D}$ are the elements on the diagonal, and the eigenvectors form the standard basis of the space $\mathbb{K}^n$.

Properties of eigenvalues and eigenvectors

Observation: Eigenvectors corresponding to the same eigenvalue form a subspace.

Proof: Consider an eigenvalue λ of a linear map f and the set $U = \{\boldsymbol{u} \in V : f(\boldsymbol{u}) = \lambda \boldsymbol{u}\}$

For any $\boldsymbol{u}, \boldsymbol{v} \in U$ we get:

- $f(a\boldsymbol{u}) = af(\boldsymbol{u}) = a\lambda\boldsymbol{u} = \lambda(a\boldsymbol{u})$,
- $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v}) = \lambda\boldsymbol{u} + \lambda\boldsymbol{v} = \lambda(\boldsymbol{u} + \boldsymbol{v})$.

Hence U is closed under addition and scalar multiples, i.e. a subspace of V.

Definition: The *geometric multiplicity* of an eigenvalue is the dimension of the space of its eigenvectors.

Properties of eigenvalues and eigenvectors

Theorem: Let $f : V \to V$ be a linear map and $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of f and $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ the corresponding nontrivial eigenvectors. Then $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ are linearly independent.

Proof: Assume for a contradiction, that k is the smallest number for which exist $\lambda_1, \dots, \lambda_k$ and $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ contradicting the claim, i.e. there are $a_1, \dots, a_k \in \mathbb{K} \setminus 0$ such that $\sum_{i=1}^{k} a_i \boldsymbol{u}_i = \mathbf{0}$.

We express $\mathbf{0}$ in two ways: $\mathbf{0} = \lambda_k \mathbf{0} = \lambda_k \sum_{i=1}^{k} a_i \boldsymbol{u}_i = \sum_{i=1}^{k} \lambda_k a_i \boldsymbol{u}_i$,

and also: $\mathbf{0} = f(\mathbf{0}) = f\left(\sum_{i=1}^{k} a_i \boldsymbol{u}_i\right) = \sum_{i=1}^{k} a_i f(\boldsymbol{u}_i) = \sum_{i=1}^{k} \lambda_i a_i \boldsymbol{u}_i$,

hence: $\mathbf{0} = \mathbf{0} - \mathbf{0} = \sum_{i=1}^{k} \lambda_i a_i \boldsymbol{u}_i - \sum_{i=1}^{k} \lambda_k a_i \boldsymbol{u}_i = \sum_{i=1}^{k-1} (\lambda_i - \lambda_k) a_i \boldsymbol{u}_i$.

As $\lambda_i \neq \lambda_k$ we get $(\lambda_i - \lambda_k) a_i \neq 0$. Already $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k-1}$ are linearly dependent — a contradiction with the minimality of k.

Properties of eigenvalues and eigenvectors

Theorem: Let $f : V \to V$ be a linear map and $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of f and $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ the corresponding nontrivial eigenvectors. Then $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ are linearly independent.

Corollary: A matrix of order n may have at most n distinct eigenvalues.

Characteristic polynomial

Definition: The *characteristic polynomial* of a matrix $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ is $p_{\boldsymbol{A}}(t) = \det(\boldsymbol{A} - t\boldsymbol{I}_n)$

Theorem: A number $\lambda \in \mathbb{K}$ is an eigenvalue of a matrix $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ if and only if λ is a root of its characteristic polynomial $p_{\boldsymbol{A}}(t)$.

Proof: λ is an eigenvalue of $\boldsymbol{A} \Leftrightarrow$

$\Leftrightarrow \exists \boldsymbol{x} \in \mathbb{K}^n \setminus \mathbf{0} : \boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$

$\Leftrightarrow \exists \boldsymbol{x} \in \mathbb{K}^n \setminus \mathbf{0} : \mathbf{0} = \boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{I}_n\boldsymbol{x} = (\boldsymbol{A} - \lambda\boldsymbol{I}_n)\boldsymbol{x}$

$\Leftrightarrow$ the matrix $\boldsymbol{A} - \lambda\boldsymbol{I}_n$ is singular

$\Leftrightarrow 0 = \det(\boldsymbol{A} - \lambda\boldsymbol{I}_n) = p_{\boldsymbol{A}}(\lambda)$

Eigenvalues — roots of the characteristic polynomial

Zero matrix:

$$\mathbf{0}_n = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \qquad p_{\mathbf{0}_n}(t) = \begin{vmatrix} -t & 0 & \dots & 0 \\ 0 & -t & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -t \end{vmatrix} = (-t)^n$$

The matrix $\mathbf{0}_n$ has only single eigenvalue 0 with multiplicity n.

A diagonal or a triangular matrix (also the identity matrix $\boldsymbol{I}_n$):

$$\boldsymbol{A} = \begin{pmatrix} a_{1,1} & * & \dots & * \\ 0 & a_{2,2} & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix} \quad p_{\boldsymbol{A}}(t) = \begin{vmatrix} a_{1,1} - t & * & \dots & * \\ 0 & a_{2,2} - t & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} - t \end{vmatrix} =$$

$$= \prod_{i=1}^{n} (a_{i,i} - t)$$

The eigenvalues of $\boldsymbol{A}$ are $a_{1,1}, a_{2,2}, \dots, a_{n,n}$.

The matrix with ones:

$$\mathbf{1}_n = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \qquad p_{\mathbf{1}_n}(t) = \begin{vmatrix} 1-t & 1 & \dots & 1 \\ 1 & 1-t & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 1-t \end{vmatrix} =$$

$$= \begin{vmatrix} -t & 0 & \dots & 0 & t \\ 0 & -t & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & -t & t \\ 1 & \dots & \dots & 1 & 1-t \end{vmatrix} = \begin{vmatrix} -t & 0 & \dots & \dots & 0 \\ 0 & -t & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -t & 0 \\ 1 & \dots & \dots & 1 & n-t \end{vmatrix} =$$

$$= (-t)^{n-1}(n-t)$$

The matrix $\mathbf{1}_n$ has the eigenvalue 0 of multiplicity $n-1$ and the eigenvalue n of multiplicity 1.

Observation: A polynomial
$(t^n + b_{n-1}t^{n-1} + \cdots + b_1 t + b_0)(-1)^{n+1}$
is the characteristic polynomial of the matrix:

$$\begin{pmatrix} 0 & 0 & \dots & 0 & -b_0 \\ 1 & 0 & \dots & 0 & -b_1 \\ 0 & 1 & \dots & 0 & -b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -b_{n-1} \end{pmatrix}$$

Proof: By expansion along the last column of:

$$\begin{vmatrix} -t & 0 & \dots & 0 & -b_0 \\ 1 & -t & \ddots & 0 & -b_1 \\ 0 & 1 & \ddots & 0 & -b_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -b_{n-1} - t \end{vmatrix}$$

(Also we can expand along the first row and reslove the recurrence.)

Computing eigenvalues and eigenvectors

Determine eigenvalues and eigenvectors of the matrix $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 3 \\ 1 & -2 & 2 \end{pmatrix}$

Characteristic polynomial: $p_{\boldsymbol{A}}(t) = \begin{vmatrix} 1-t & 2 & 0 \\ 3 & -1-t & 3 \\ 1 & -2 & 2-t \end{vmatrix} =$

$= (1-t)(-1-t)(2-t) + 6 + 6(1-t) - 6(2-t) = -t^3 + 2t^2 + t - 2$

The eigenvalues of $\boldsymbol{A}$ are the roots of $p_{\boldsymbol{A}}(t)$, i.e. 2, 1 and -1.

The eigenvector $\boldsymbol{x}_1$ for $\lambda_1 = 2$ is any solution of $(\boldsymbol{A} - \lambda_1 \boldsymbol{I}_3)\boldsymbol{x}_1 = \mathbf{0}$

$$\boldsymbol{A} - \lambda_1 \boldsymbol{I}_3 = \begin{pmatrix} 1-2 & 2 & 0 \\ 3 & -1-2 & 3 \\ 1 & -2 & 2-2 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 0 \\ 3 & -3 & 3 \\ 1 & -2 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

The solution $\boldsymbol{x}_1$ is any scalar multiple of the vector $(2, 1, -1)^T$.

The eigenvalue $\lambda_2 = 1$ yields the eigenvector $\boldsymbol{x}_2 = (-1, 0, 1)^T$, and the eigenvalue $\lambda_3 = -1$ yields the eigenvector $\boldsymbol{x}_3 = (-1, 1, 1)^T$.

Coefficients of the characteristic polynomial

Observation: For $p_{\boldsymbol{A}}(t) = \det(\boldsymbol{A} - t\boldsymbol{I}_n) = \sum_{i=0}^{n} b_i t^i$ it holds:

- $b_n = (-1)^n$ ... only the product along the diagonal in $\boldsymbol{A} - t\boldsymbol{I}_n$ may yield t^n, each its factor of t has coefficient -1.
- $b_0 = \det(\boldsymbol{A})$ ... substitute $t = 0$ into $p_{\boldsymbol{A}}(t)$
- $b_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{i,i}$

... the term t^{n-1} could be obtained only from the product of t linear terms $a_{i,i} - t$ that are on the diagonal of $\boldsymbol{A} - t\boldsymbol{I}_n$ by choosing $n-1$ times the term $-t$ and once each of $a_{i,i}$. There are n choices, where summands $a_{i,i}$ in the coefficient $b_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{i,i}$ correspond to distinct choices.

Coefficients of the characteristic polynomial

Observation: For $p_{\boldsymbol{A}}(t) = \det(\boldsymbol{A} - t\boldsymbol{I}_n) = \sum_{i=0}^{n} b_i t^i$ it holds:

- $b_n = (-1)^n$ ... only the product along the diagonal in $\boldsymbol{A} - t\boldsymbol{I}_n$ may yield t^n, each its factor of t has coefficient -1.
- $b_0 = \det(\boldsymbol{A})$ ... substitute $t = 0$ into $p_{\boldsymbol{A}}(t)$
- $b_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{i,i}$

Review: When $\mathbb{K}$ is algebraically closed, one may factorize the characteristic polynomial into linear factors with roots/eigenvalues: $p_{\boldsymbol{A}}(t) = (\lambda_1 - t)^{r_1}(\lambda_2 - t)^{r_2} \ldots (\lambda_k - t)^{r_k}$ with $r_1 + \cdots + r_k = n$. The exponent r_i is the *algebraic multiplicity* of the eigenvalue λ_i.

Coefficients of the characteristic polynomial

Observation: For $p_{\boldsymbol{A}}(t) = \det(\boldsymbol{A} - t\boldsymbol{I}_n) = \sum_{i=0}^{n} b_i t^i$ it holds:

- $b_n = (-1)^n$ ... only the product along the diagonal in $\boldsymbol{A} - t\boldsymbol{I}_n$ may yield t^n, each its factor of t has coefficient -1.
- $b_0 = \det(\boldsymbol{A})$ ... substitute $t = 0$ into $p_{\boldsymbol{A}}(t)$
- $b_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{i,i}$

Moreover, if the characteristic polynomial has a decomposition $p_{\boldsymbol{A}}(t) = (\lambda_1 - t)^{r_1}(\lambda_2 - t)^{r_2} \dots (\lambda_k - t)^{r_k}$, $r_1 + \cdots + r_k = n$ then:

- $b_0 = \det(\boldsymbol{A}) = \prod_{i=1}^{k} \lambda_i^{r_i}$ ... again substitute $t = 0$ into $p_{\boldsymbol{A}}(t)$
- $b_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{i,i} = (-1)^{n-1} \sum_{i=1}^{k} \underbrace{\lambda_i + \cdots + \lambda_i}_{r_i}$

  ... analogously from $(\lambda_1 - t)^{r_1}(\lambda_2 - t)^{r_2} \dots (\lambda_k - t)^{r_k}$.

Application — eigenvalues of rabbit populations

Assume that the rabbits' breeding is described by a simple law, e.g. that this year's number of rabbits is the sum of these numbers of the past two years.

Let F_t denotes the number of rabbits in year t. We get the recurrent formula for Fibbonacci numbers $F_t = F_{t-1} + F_{t-2}$.
We may ask, how the ratio $\frac{F_t}{F_{t-1}}$ develops — whether it has a limit, or whether it oscillates, or whether it becomes stable.

The same in the language of matrices and vector spaces:
Consider the space $\mathbb{R}^2$, then the linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\begin{pmatrix} F_t \\ F_{t-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{t-1} \\ F_{t-2} \end{pmatrix}$$

describes the same recurrent relation.
E.g., if we start with a single rabbit, we get the sequence

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \xrightarrow{f} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{f} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \xrightarrow{f} \begin{pmatrix} 3 \\ 2 \end{pmatrix} \xrightarrow{f} \begin{pmatrix} 5 \\ 3 \end{pmatrix} \xrightarrow{f} \begin{pmatrix} 8 \\ 5 \end{pmatrix} \xrightarrow{f} \begin{pmatrix} 13 \\ 8 \end{pmatrix} \xrightarrow{f} \dots$$

The stable ratio $\frac{F_t}{F_{t-1}}$, have vectors $\boldsymbol{x} = \begin{pmatrix} F_t \\ F_{t-1} \end{pmatrix}$ satisfying:

$$f(\boldsymbol{x}) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \boldsymbol{x} = \lambda \boldsymbol{x}$$

for some $\lambda \in \mathbb{R}$. (Vectors $\boldsymbol{x}$ and $\lambda\boldsymbol{x}$ have the same ratio.)

The matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ has two eigenvalues, namely:

$\lambda_1 = \frac{1+\sqrt{5}}{2}$ $\quad \boldsymbol{x} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix}$ . . . the vector's elements grow with every iteration

$\lambda_2 = \frac{1-\sqrt{5}}{2}$ $\quad \boldsymbol{x} = \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}$ . . . here the sign changes with every iteration, and the limit is the zero vector

Solving a homogeneous system of first order linear differential equations with constant coefficients

$$\begin{array}{ccccccc} y_1' & = & a_{1,1}y_1 & + & \dots & + & a_{1,n}y_n \\ \vdots & & \vdots & & & & \vdots \\ y_n' & = & a_{n,1}y_1 & + & \dots & + & a_{n,n}y_n \end{array} \quad \text{yield } \boldsymbol{A} = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$$

For an eigenvalue λ and an eigenvector $(k_1, \dots, k_n)^T$ of $\boldsymbol{A}$, the n-tuple of functions $y_i(x) := k_i e^{\lambda x}$ solves the original system:

$$\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} (k_1 e^{\lambda x})' \\ \vdots \\ (k_n e^{\lambda x})' \end{pmatrix} = \begin{pmatrix} \lambda k_1 e^{\lambda x} \\ \vdots \\ \lambda k_n e^{\lambda x} \end{pmatrix} = e^{\lambda x} \lambda \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = e^{\lambda x} \boldsymbol{A} \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} =$$

$$= \begin{pmatrix} a_{1,1}k_1 e^{\lambda x} & + \dots + & a_{1,n}k_n e^{\lambda x} \\ \vdots & & \vdots \\ a_{n,1}k_1 e^{\lambda x} & + \dots + & a_{n,n}k_n e^{\lambda x} \end{pmatrix} = \begin{pmatrix} a_{1,1}y_1 & + \dots + & a_{1,n}y_n \\ \vdots & & \vdots \\ a_{n,1}y_1 & + \dots + & a_{n,n}y_n \end{pmatrix}$$

Note that other eigenvalues yield another sets of particular solutions.

The overall solution is any linear combination of particular solutions that satisfies all boundary conditions.