

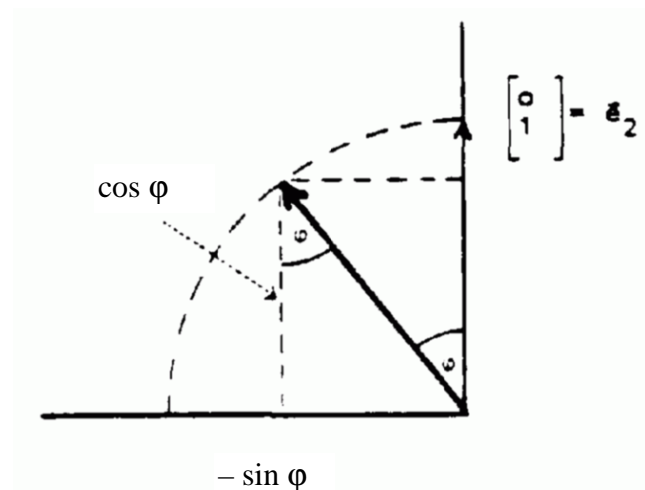
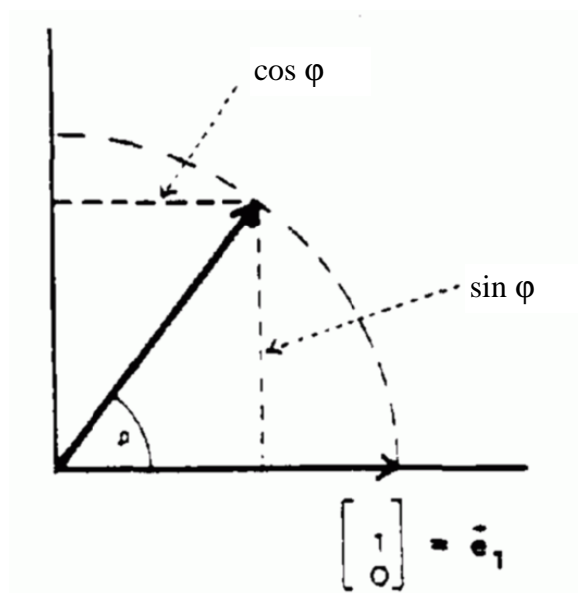
Special cases of linear mappings

(a) Rotations around the origin

Let f be the counterclockwise rotation by the angle φ around the zero point $(0; 0)$ (origin of the cartesian coordinate system).

Each vector is rotated by φ , its image has the same length as before.

Image vectors of the standard basis vectors = ?



We obtain as image vectors: $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$ and $\begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$.

The matrix of f is thus:

$$A = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

We call this a *rotation matrix*.

(b) Scaling

Let f be the mapping which enlarges (or shrinks) every vector by a certain, fixed factor $\lambda \neq 0$.

Its corresponding matrix is

$$A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix},$$

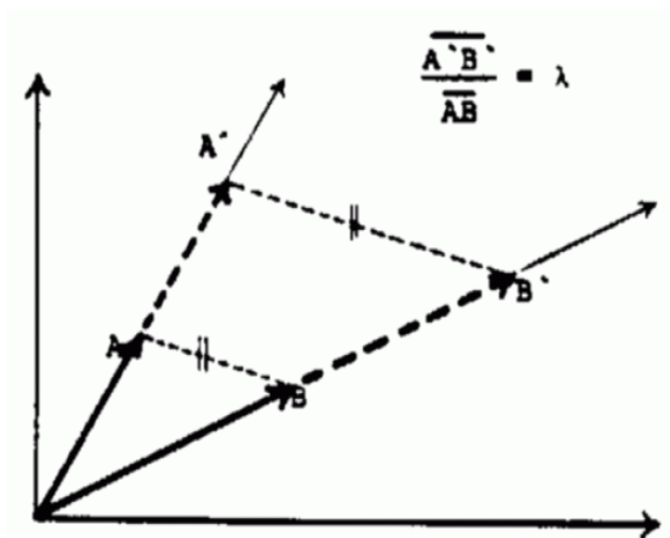
called a *scaling matrix* with factor λ .

Indeed, we have

$$A \vec{x} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix}$$

The image of each vector has the same direction as before (or the opposite direction, if λ is negative), but a length which is modified by the factor $|\lambda|$.

Parallelism and all angles remain unchanged under this mapping.



(c) Centraffine mapping

Let f act as a scaling by λ_1 on the x axis and as a scaling by λ_2 on the y axis, with two different numbers $\lambda_1 \neq \lambda_2$.

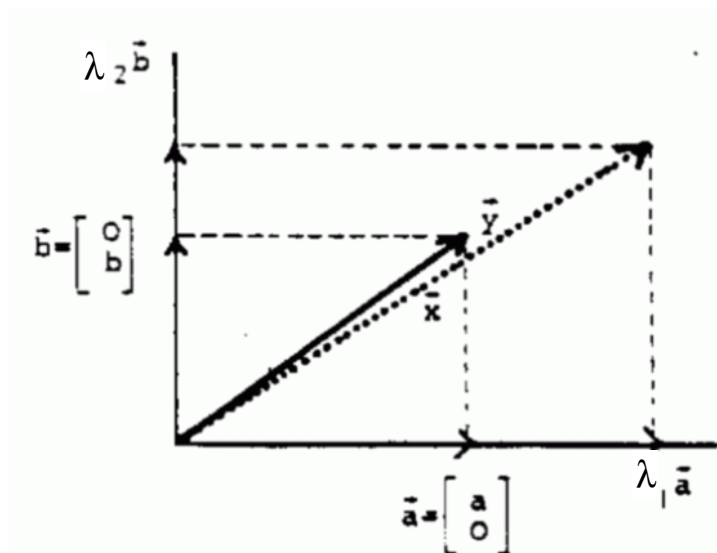
The corresponding matrix is

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

This mapping is called *centraffine*.

Its effect: $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \lambda_1 a \\ \lambda_2 b \end{bmatrix}$

It works as a pure scaling on the x and y axis, but not for vectors which are outside these coordinate axes (they are also rotated a bit):



The centraffine mapping is thus not a scaling for all vectors.

Certain vectors play a special role for this mapping, namely, those on the coordinate axes: They are only scaled, the others are also rotated.

11. Eigenvalues and eigenvectors

We have seen in the last chapter:
for the centroaffine mapping

$$f: \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \end{bmatrix},$$

some directions, namely, the directions of the coordinate axes: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, are distinguished among all directions in the plane: In them, f acts as a *pure scaling*.

We want to generalize this to arbitrary linear mappings.

We call a vector representing such a direction an *eigenvector* of the linear mapping f (or of the corresponding matrix A), and the scaling factor which describes the effect of f on it an *eigenvalue*.

Examples:

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is eigenvector of the matrix $\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$ to the eigenvalue 3:

$$\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} = 3 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ is also eigenvector of $\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$ to the eigenvalue 3:

$$\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} = 3 \cdot \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is eigenvector of $\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$ to the eigenvalue 7:

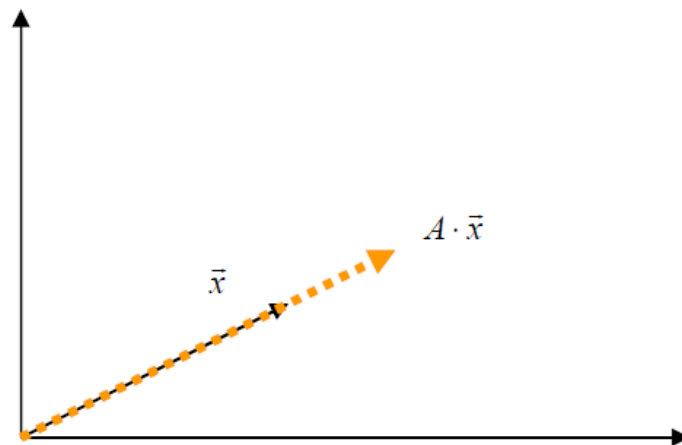
$$\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \end{bmatrix} = 7 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

in general:

An eigenvector of A must fulfill $A \cdot \vec{x} = \lambda \cdot \vec{x}$, and we require $\vec{x} \neq \vec{0}$.

Definition:

Let A be a matrix of type (n, n) . If there exists a real number λ such that the equation $A \cdot \vec{x} = \lambda \cdot \vec{x}$ has a solution $\vec{x}_\lambda \neq \vec{0}$, we call λ an *eigenvalue* and \vec{x}_λ an *eigenvector* of the matrix A .



If \vec{x}_λ is an eigenvector of A and $a \neq 0$ an arbitrary factor, then also $a \cdot \vec{x}_\lambda$ is an eigenvector of A .

We can choose a in a way that the length of $a \cdot \vec{x}_\lambda$ becomes 1. That means, we can always find eigenvectors of length 1.

If we insert $\vec{x} = E \vec{x}$, we can transform the equation $A \cdot \vec{x} = \lambda \cdot \vec{x}$ in the following way:

$$\begin{aligned} A \cdot \vec{x} = \lambda \cdot \vec{x} &\Leftrightarrow A \cdot \vec{x} - \lambda \cdot \vec{x} = \vec{0} \\ &\Leftrightarrow A \vec{x} - \lambda E \vec{x} = \vec{0} \\ &\Leftrightarrow (A - \lambda E) \vec{x} = \vec{0} \end{aligned}$$

This is equivalent to a system of linear equations with matrix $A - \lambda E$ and with right-hand side always zero.

If the matrix $A - \lambda E$ has maximal rank (i.e., if it is regular), this system has exactly one solution (i.e., the trivial solution: the zero vector). We are *not* interested in *that* solution!

The system has other solutions (infinitely many ones), if and only if $A - \lambda E$ is singular, that means, if and only if

$$\det(A - \lambda E) = 0.$$

From this, we can derive a method to determine all eigenvalues and eigenvectors of a given matrix.

The equation $\det(A - \lambda E) = 0$ (called the *characteristic equation* of A) is an equation between numbers (not vectors) and includes the unknown λ . Solving it for λ means finding all possible eigenvalues of A .

In the case of a 2×2 matrix A , the characteristic equation $\det(A - \lambda E) = 0$ has the form

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0$$

i.e., it is a quadratic equation and can be solved with the well-known pq formula (see Chapter 6, p. 28).

Example:

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

$$A - \lambda E = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} - \lambda \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \lambda & -\frac{1}{2} \\ -\frac{1}{2} & 1 - \lambda \end{bmatrix}$$

$$\det(A - \lambda E) = \begin{vmatrix} 1 - \lambda & -\frac{1}{2} \\ -\frac{1}{2} & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 - \left(-\frac{1}{2}\right)^2 = 1 - 2\lambda + \lambda^2 - \frac{1}{4}$$

$$= \lambda^2 - 2\lambda + \frac{3}{4} \stackrel{!}{=} 0$$

$$\Leftrightarrow \lambda_{1,2} = 1 \pm \sqrt{1 - \frac{3}{4}} = 1 \pm \frac{1}{2}, \quad \lambda_1 = \frac{1}{2}, \quad \lambda_2 = \frac{3}{2}$$

$\lambda^2 - 2\lambda + \frac{3}{4}$ is called the *characteristic polynomial* of A .

Its zeros, the solutions $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{3}{2}$, are the eigenvalues of A .

That means: Exactly for $\lambda = \frac{1}{2}$ and $\lambda = \frac{3}{2}$ does the vector equation $A \cdot \vec{x} = \lambda \cdot \vec{x}$ have nontrivial solution vectors $\vec{x} \neq \vec{0}$, i.e., eigenvectors.

The next step is to find these eigenvectors for each of the eigenvalues:

This means to solve a system of linear equations!

We use the equivalent form $(A - \lambda E) \vec{x} = \vec{0}$.

We are not interested in the trivial solution $\vec{x} = \vec{0}$.

In the example: To find an eigenvector

to the eigenvalue $\lambda_1 = \frac{1}{2}$: $(A - \frac{1}{2}E) \vec{x} = \vec{0}$

$$\Leftrightarrow \begin{bmatrix} 1 - \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 - \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Leftrightarrow \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(system of 2 linear equations with r.h.s. 0)

with elementary row operations we get:

$$\begin{array}{cc|c} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ \hline 1 & -1 & 0 \\ 1 & -1 & 0 \\ \hline 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}$$

From the second-last row we deduce:

$$x_1 + (-x_2) = 0$$

We can choose one parameter arbitrarily, e.g., $x_2 = c$, and obtain the general solution

$\vec{x} = \begin{bmatrix} c \\ c \end{bmatrix}$ (with $c \in \mathbb{R}$ and $c \neq 0$ because we want to have an eigenvector)

It is enough to give just one vector as a representative of this direction, e.g.,

$$\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This is an eigenvector of A to the eigenvalue $1/2$.

$$\text{Test: } A \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda_1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenvectors to the second eigenvalue, $3/2$, are determined analogously

(a solution is $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.)

In the general case of an $n \times n$ matrix, $\det(A - \lambda E)$ is a *polynomial* in the variable λ of degree n , i.e., when we develop the determinant, we get something of the form

$$c_n \lambda^n + c_{n-1} \lambda^{n-1} + \dots + c_1 \lambda + c_0$$

Such a polynomial has at most n zeros, so A can have at most n different eigenvalues.

Attention:

There are matrices which have no (real) eigenvalues at all!

Example: Rotation matrices with angle $\varphi \neq 0^\circ, 180^\circ$.

It is also possible that for the same eigenvalue, there are different eigenvectors with different directions.

Example: For the scaling matrix $A = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$, every vector $\vec{x} \neq \vec{0}$ is eigenvector to the eigenvalue 5.

Fixed points and attractors

Let $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be an arbitrary mapping.

$\vec{x} \in \mathbb{R}^n$ is called a **fixed point** of f , if $f(\vec{x}) = \vec{x}$, i.e., if \vec{x} remains "fixed" under the mapping f .

\vec{x} is called *attracting fixed point*, *point attractor* or *vortex point* of f , if there exists additionally a neighbourhood of \vec{x} such that for each \vec{y} from this neighbourhood the sequence

$$\vec{y}, f(\vec{y}), f(f(\vec{y})), \dots$$

converges against \vec{x} .

The fixed points of linear mappings are exactly (by definition) the *eigenvectors to the eigenvalue 1* and the zero vector.

Examples:

$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ (shear mapping): each point on the x axis is a fixed point.

$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ (scaling by 2): only the origin (0; 0) is fixed point. (There are no eigenvectors to the eigenvalue 1; the only eigenvalue is 2.)
The origin is not attracting.

$A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ (scaling by 1/2, i.e., shrinking):
the origin (0; 0) is attracting fixed point.

Definition:

A *stochastic matrix* is an $n \times n$ matrix where all columns sum up to 1.

Theorem:

Each stochastic matrix has the eigenvalue 1.

The corresponding linear mapping has thus a fixed point $\neq \vec{0}$.

Example from epidemiology:

The outbreak of a disease is conceived as a stochastic (random) process. For a tree there are two possible states:

"healthy" (state 0) and
"infected" (state 1).

For a healthy tree, let us assume a probability of 1/4 to be infected after one year, i.e.:

$p_{01} = \frac{1}{4}$, and correspondingly: $p_{00} = \frac{3}{4}$ (= probability to stay healthy).

For infected trees, we assume a probability of spontaneous recovery of 1/3:

$p_{10} = \frac{1}{3}$, $p_{11} = \frac{2}{3}$

We define the *transition matrix* (similar to the age-classes example) as

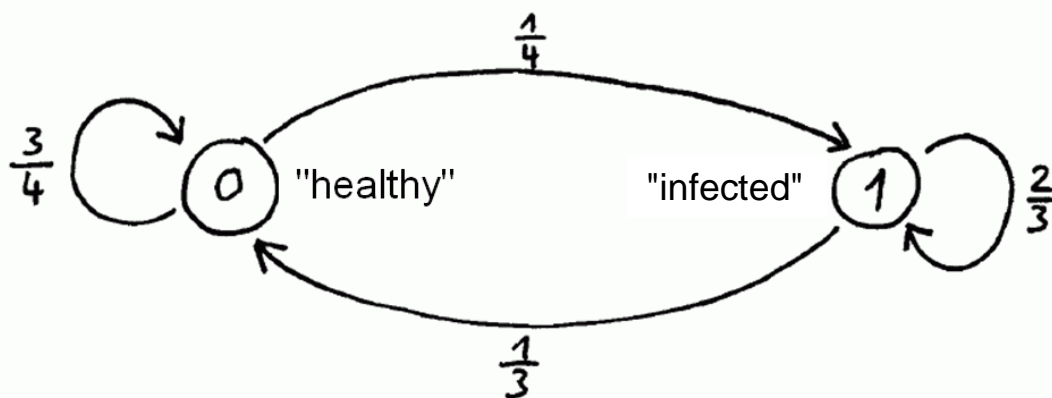
$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

For the purpose of calculation, we need the transposed of P , which is a stochastic matrix (and is in the literature also often called the transition matrix):

$$P^T = \begin{bmatrix} \frac{3}{4} & \frac{1}{3} \\ \frac{1}{4} & \frac{2}{3} \end{bmatrix}$$

A process of this sort, where the probability to come into a new state depends only on the current state, is called a *Markov chain*.

Graphical representation of the transitions:



If we assume that g_1 , resp., k_1 are the proportions of healthy, resp., infected trees in the first year, the average proportions in the 2nd year are given by:

$$\begin{bmatrix} g_2 \\ k_2 \end{bmatrix} = P^T \cdot \begin{bmatrix} g_1 \\ k_1 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \cdot g_1 + \frac{1}{3} \cdot k_1 \\ \frac{1}{4} \cdot g_1 + \frac{2}{3} \cdot k_1 \end{bmatrix}$$

Question: what is the percentage of infected trees, if the tree stand is undisturbed for many years and the transition probabilities remain the same?

We have to look for a fixed point of the mapping corresponding to P^T .

Because P^T is a stochastic matrix, it has automatically the eigenvalue 1.

We have only to determine a corresponding eigenvector (fixed point) $\begin{pmatrix} g' \\ k' \end{pmatrix}$:

$$\begin{bmatrix} \frac{3}{4}-1 & \frac{1}{3} \\ \frac{1}{4} & \frac{2}{3}-1 \end{bmatrix} \cdot \begin{bmatrix} g' \\ k' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

By applying the standard method for solving linear systems, we obtain:

$$\begin{bmatrix} g' \\ k' \end{bmatrix} = \begin{bmatrix} 4 \cdot c \\ 3 \cdot c \end{bmatrix}, c \neq 0$$

From this we derive the proportion of the infected trees:

$$\frac{k}{g+k} = \frac{3}{4+3} = \frac{3}{7}$$

Remarks:

This proportion does not depend on the number of infected trees in the first year.

$\begin{pmatrix} g' \\ k' \end{pmatrix}$ is in fact an attracting fixed point, if we restrict ourselves to a fixed total number of trees, $g+k$.

In the same way, a *stable age-class distribution* can be calculated in the case of the age-class transition matrix (see Chapter 10, p. 82-83).

In that case, the stable age-class vector \vec{a}^* has to be determined as the fixed point (eigenvector to the eigenvalue 1) of the matrix P^T , i.e., as the solution to

$$P^T \cdot \vec{a}^* = \vec{a}^* .$$

Because the fixed point is attracting, it can be obtained as the limit of the sequence

\vec{a}_0 , $P^T \cdot \vec{a}_0$, $(P^T)^2 \cdot \vec{a}_0$, $(P^T)^3 \cdot \vec{a}_0$, ... ,
starting from an initial vector \vec{a}_0 .