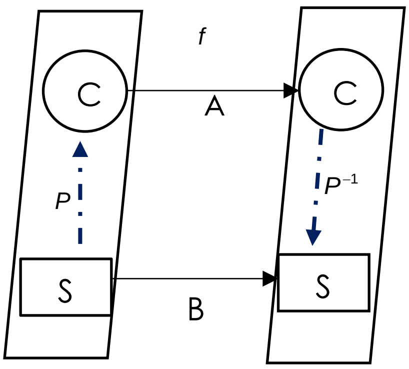
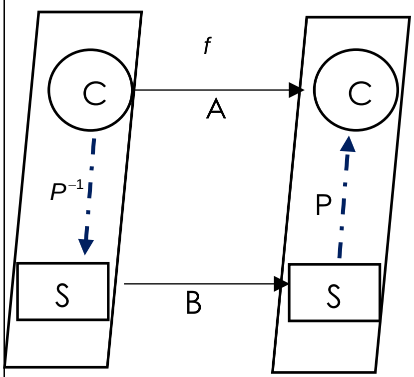
2. Diagonalization

2.1 Eigenvalues and Eigenvectors

2.1.2 Introduction and definitions

Definition

A matrix A is similar to a matrix B if there is an invertible matrix P, such that $A = P^{-1}BP$

$P^{-1}AP = B$	$A = PBP^{-1}$
	

Definition 4.6. Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a square matrix. Then $\lambda \in \mathbb{R}$ is an *eigenvalue* of $\boldsymbol{A}$ and $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$ is the corresponding *eigenvector* of $\boldsymbol{A}$ if

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}\,. \tag{4.25}$$

- If u is eigenvector associated with λ then ku is also eigenvector associated with λ.
- If u, v are linearly independent eigenvectors associated with λ then $au + bv$ is also eigenvector associated with λ.
- There can be linearly independent eigenvectors associated with the same eigenvalue.
- The scalar 0 can be an endomorphism's eigenvalue. In this case, any eigenvector associated with 0 belong to the kernel, therefore, the endomorphism cannot be injective.

2.1.3 Eigenvalues and Spectrum

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}\begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \lambda \begin{bmatrix} \\ x \\ \\ \end{bmatrix} \Leftrightarrow \lambda \begin{bmatrix} \\ x \\ \\ \end{bmatrix} - \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}\begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Leftrightarrow$$

$$\left(\lambda \begin{bmatrix} & & \\ & I_n & \\ & & \end{bmatrix} - \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}\right)\begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

- λ is an eigenvalue of $\boldsymbol{A} \in \mathbb{R}^{n\times n}$.
- There exists an $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$ with $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, or equivalently, $(\boldsymbol{A} - \lambda\boldsymbol{I}_n)\boldsymbol{x} = \boldsymbol{0}$ can be solved non-trivially, i.e., $\boldsymbol{x} \neq \boldsymbol{0}$.
- $\mathrm{rk}(\boldsymbol{A} - \lambda\boldsymbol{I}_n) < n$.
- $\det(\boldsymbol{A} - \lambda\boldsymbol{I}_n) = 0$.

Definition 4.5 (Characteristic Polynomial)**.** For $\lambda \in \mathbb{R}$ and a square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$

$$p_{\boldsymbol{A}}(\lambda) := \det(\boldsymbol{A} - \lambda\boldsymbol{I}) \tag{4.22a}$$

Theorem 4.8. *$\lambda \in \mathbb{R}$ is eigenvalue of $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ if and only if λ is a root of the characteristic polynomial $p_{\boldsymbol{A}}(\lambda)$ of $\boldsymbol{A}$.*

Definition 4.9. Let a square matrix $\boldsymbol{A}$ have an eigenvalue λ_i. The *algebraic multiplicity* of λ_i is the number of times the root appears in the characteristic polynomial.

Definition 4.10 (Eigenspace and Eigenspectrum)**.** For $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, the set of all eigenvectors of $\boldsymbol{A}$ associated with an eigenvalue λ spans a subspace of $\mathbb{R}^n$, which is called the *eigenspace* of $\boldsymbol{A}$ with respect to λ and is denoted by E_λ. The set of all eigenvalues of $\boldsymbol{A}$ is called the *eigenspectrum*, or just *spectrum*, of $\boldsymbol{A}$.

- Similar matrices (see Definition 2.22) possess the same eigenvalues. Therefore, a linear mapping Φ has eigenvalues that are independent of the choice of basis of its transformation matrix. This makes eigenvalues, together with the determinant and the trace, key characteristic parameters of a linear mapping as they are all invariant under basis change.
- Symmetric, positive definite matrices always have positive, real eigenvalues.

2.1.4 Eigenvectors and eigenspace

Suppose that x is a eigenvector associate to the eigenvalue λ, then

$$Ax = \lambda x$$

$$Ax = \lambda x \Leftrightarrow Ax - \lambda x = 0_{n\times 1} \Leftrightarrow (A - \lambda I_n)x = 0_{n\times 1} \, .$$

Definition 4.11. Let λ_i be an eigenvalue of a square matrix $\boldsymbol{A}$. Then the *geometric multiplicity* of λ_i is the number of linearly independent eigenvectors associated with λ_i. In other words, it is the dimensionality of the eigenspace spanned by the eigenvectors associated with λ_i.

Remark. A specific eigenvalue's geometric multiplicity must be at least one because every eigenvalue has at least one associated eigenvector. An eigenvalue's geometric multiplicity cannot exceed its algebraic multiplicity, but it may be lower. $\diamondsuit$

Graphical Intuition in Two Dimensions

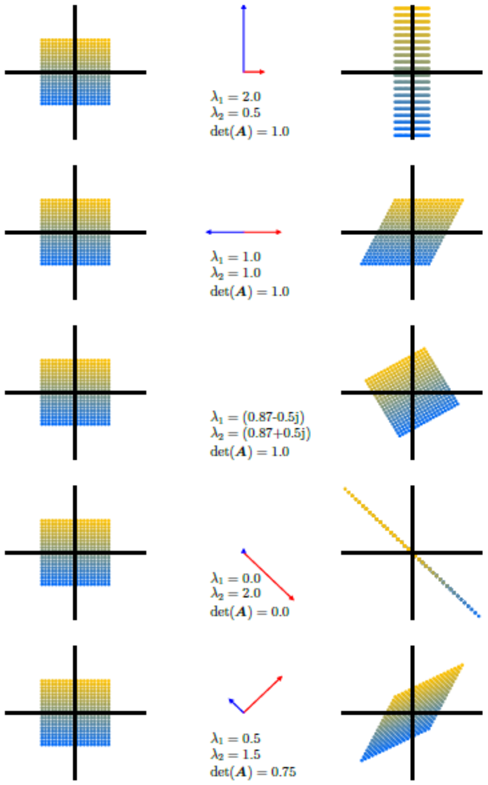


Figure 4.4 Determinants and eigenspaces. Overview of five linear mappings and their associated transformation matrices $\boldsymbol{A}_i \in \mathbb{R}^{2\times 2}$ projecting 400 color-coded points $\boldsymbol{x} \in \mathbb{R}^2$ (left column) onto target points $\boldsymbol{A}_i\boldsymbol{x}$ (right column). The central column depicts the first eigenvector, stretched by its associated eigenvalue λ_1, and the second eigenvector stretched by its eigenvalue λ_2. Each row depicts the effect of one of five transformation matrices $\boldsymbol{A}_i$ with respect to the standard basis .

2.2 Eigendecomposition and Diagonalization

Definition 4.19 (Diagonalizable). A matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is *diagonalizable* if it is similar to a diagonal matrix, i.e., if there exists an invertible matrix $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ such that $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

Theorem 4.20 (Eigendecomposition). *A square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ can be factored into*

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}\,, \tag{4.55}$$

where $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{D}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\boldsymbol{A}$, if and only if the eigenvectors of $\boldsymbol{A}$ form a basis of $\mathbb{R}^n$.

Theorem 4.12. *The eigenvectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ with n distinct eigenvalues $\lambda_1, \dots, \lambda_n$ are linearly independent.*

Theorem

The eigenvectors of a matrix $A \in \mathbb{R}^{n\times n}$ form a basis of $\mathbb{R}^n$ if and only if for distinct eigenvalues λ_1, λ_2, ..., $\lambda_k \in \mathbb{R}$, we have:

- $ma(\lambda_i) = mg(\lambda_i)$, with $i = 1, \dots, k$
- $n = mg(\lambda_1) + mg(\lambda_2) + \dots + mg(\lambda_k)$

Theorem 4.15 (Spectral Theorem). *If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is symmetric, there exists an orthonormal basis of the corresponding vector space V consisting of eigenvectors of $\boldsymbol{A}$, and each eigenvalue is real.*

Theorem 4.21. *A symmetric matrix $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ can always be diagonalized.*

Theorem 4.16. *The determinant of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is the product of its eigenvalues, i.e.,*

$$\det(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i\,, \tag{4.42}$$

where λ_i are (possibly repeated) eigenvalues of $\boldsymbol{A}$.

Theorem 4.17. *The trace of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is the sum of its eigenvalues, i.e.,*

$$tr(\boldsymbol{A}) = \sum_{i=1}^{n} \lambda_i\,, \tag{4.43}$$

where λ_i are (possibly repeated) eigenvalues of $\boldsymbol{A}$.

- Diagonal matrices $\boldsymbol{D}$ can efficiently be raised to a power. Therefore, we can find a matrix power for a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ via the eigenvalue decomposition (if it exists) so that

$$\boldsymbol{A}^k = (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^k = \boldsymbol{P}\boldsymbol{D}^k\boldsymbol{P}^{-1}\,. \tag{4.62}$$

Computing $\boldsymbol{D}^k$ is efficient because we apply this operation individually to any diagonal element.

2.4 Applications

2.4.1 Populational Growth (Poole)

We intend to study the growth evolution of a certain animal species, whose maximum age of a female is 3 years, and that the population can be divided into three age groups: young (up to 1 year); young adults (from 1 to 2 years); adults (from 2 to 3 years).

It is known that in each year the young do not reproduce, the young adults produce 3 offspring and the adults only one.

It is also known that, the probability of a young reaching young adult is 60%, while the probability of a young adult reaching adult is 40%.

If a population has j young individuals, young adult individuals, and adult individuals. What will be the predicted number of adults after one year, two years, and 10 years?

After a year the number of individuals can be estimated by:

- young individuals will be $0 \times j + 3 \times aj + 1 \times a$;
- young adult individuals will be $0.6 \times j$;
- adult individuals will be $0.4 \times aj$.

In matrix terms these calculations can be represented by: $\underbrace{\begin{bmatrix} 0 & 3 & 1 \\ 0.6 & 0 & 0 \\ 0 & 0.4 & 0 \end{bmatrix}}_{L} \begin{bmatrix} j \\ aj \\ a \end{bmatrix}$

Let $x_0 = \begin{bmatrix} j \\ aj \\ a \end{bmatrix}$ the vector with the initial population distribution and the population distribution in the year the population distribution at the end of the:

- the first year will be $x_1 = Lx_0$;
- the second year will be $x_2 = Lx_1 = L(Lx_0) = L^2 x_0$;
- the third year will be $x_3 = Lx_2 = L(L^2 x_0) = L^3 x_0$;

  …

- i-ésimo year will be $x_i = Lx_{i-1} = L(L^{i-1} x_0) = L^i x_0$.

The matrix L is called the Leslie matrix, and is used in biology to estimate population growth with n age classes of equal length. In general, the Leslie matrix is defined by

$$L=\begin{bmatrix} b_1 & b_2 & \dots & b_{n-1} & b_n \\ s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & s_{n-1} & \end{bmatrix} \in M_{n\times n}(\mathbb{R}).$$

Where is the average number of females produced by the class and is the probability that a female in the class survives (i.e. in the next stage is in the class). The population distribution at the end of the year will be calculated by $x_k = L^k x_0$.

Example: In an initial population of 40 youth, 30 young adults, and 20 adults, determine the distribution of the population after 150 years

1) Determine the number of individuals after 150 years.

2) Determine the ratio for each adult

3) Determine the growth rate

4) Determine the eigenvalues and eigenvectors of L

2.4.2 Population Predictions (Williams)

A village with a population of 4725 people has two supermarkets: In the current month 2300 of its residents usually go to supermarket A, and the remaining 2425 are customers of supermarket B. The probability that at the end of each month, a customer from supermarket A will switch to supermarket B is 10%, while 85% of the customers of supermarket B, do not change their preferences.

The Markov Matrix $P=\begin{array}{cc} \begin{array}{cc} \text{sup } A & \text{sup B} \end{array} & \\ \begin{bmatrix} 0.9 & 0.15 \\ 0.1 & 0.85 \end{bmatrix} & \begin{array}{l} \text{sup } A \\ \text{sup B} \end{array} \end{array}$ and the initial vector $x_0=\begin{bmatrix} 2300 \\ 2425 \end{bmatrix}$, allow you to study the evolution of customers over the months:

- the first year of $x_1 = Px_0$.
- at the end of the second month of $x_2 = Px_1 = P(Px_0) = P^2 x_0$.

…

- at the end of the i-ésimo month of $x_i = Px_{i-1} = P\left(P^{i-1}x_0\right) = P^i x_0$.

Example: Determine the market evolution after 100 months

1) Determine the market evolution after 100 months

2) Determine the market evolution after 101 months

3) Determine the eigenvalues and eigenvectors of P

4) Determine P^{200}

2.4.3 Ranking sports teams (Poole)

After 6 games in a snooker championship with 4 players, the results were recorded in a matrix A, as follows: $a_{ij} = 1$, if the player i win to player j, and has $a_{ij} = 0$ otherwise.

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

We would like to associate a ranking r_i with player i such a way that, $r_i > r_j$ indicates that player i is ranked more highly than player j.

For this purpose, let's require that the *ri*'s be probabilities (that is $0 \le r_i \le 1$, for all I and $r_1 + r_2 + r_3 + r_4 = 1$) and then organize the rankings in a *ranking vector*

$$r = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix}.$$

The ranking of each player is intended to be proportional to the sum of the rankings of the players defeated by player i. Let p, be the constant of proportionality:

- $r_1 = p(r_2 + r_3)$
- $r_2 = p(r_3)$
- $r_3 = p(r_4)$
- $r_4 = p(r_1 + r_2)$

Observe that we can write this system in matrix form as

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = p \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix}.$$

Equivalently, we see that the ranking vector r must satisfy, $Ar = \frac{1}{p} r$. Then $\frac{1}{p}$ is an eigenvalue and r is an eigenvector.

- Determine the eigenvalues and eigenvectors

- Determine an eigenvector such that $r = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix}$, such that $r_1 + r_2 + r_3 + r_4 = 1$.