Basics of Linear Algebra

Florian Wickelmaier

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



Summer 2026

Basics of Linear Algebra

Multivariate observations

- Based on a sample of persons (e. g. of size n), we observe realizations of p random variables $X_1, \ldots, X_p$
- This results in $n \cdot p$ values

$$x_{ij}$$

characterizing the outcome of person P_i $(i = 1, \ldots, n)$ with respect to the random variables X_j $(j = 1, \ldots, p)$

Data Scheme

The basic data scheme looks like this

		Random Variables				
		X_1	X_2	X_3	$\dots$	X_p
Persons	P_1	x_{11}	x_{12}	x_{13}	$\dots$	x_{1p}
	P_2	x_{21}	x_{22}	x_{23}	$\dots$	x_{2p}
	P_3	x_{31}	x_{32}	x_{33}	$\dots$	x_{3p}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	P_n	x_{n1}	x_{n2}	x_{n3}	$\dots$	x_{np}

Example: Multivariate Data

- Values for three persons P_1, P_2 and P_3 in two variables X_1 and X_2, representing the results of two different diagnostic tests

	X_1	X_2
P_1	1	-2
P_2	2	1
P_3	-3	1

- There are two different ways to geometrically illustrate the data

Variable Space

For each person we consider the pair of values resulting from the two tests

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \mathbf{p}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{p}_3 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

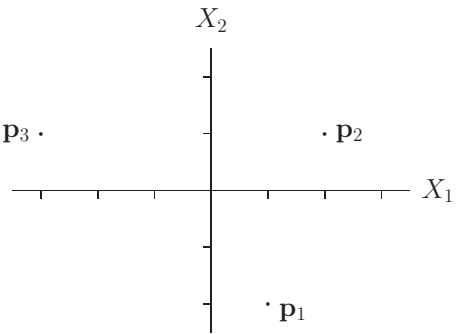
In general, this defines a *variable space*

- The axes of the coordinate system correspond to the considered variables
  - In the example results a two-dimensional space with the axes corresponding to the variables X_1 and X_2
- In this space, each person is represented by a point (i. e., by a vector)
- Graphically, this representation is a scatter plot known from introductory statistics

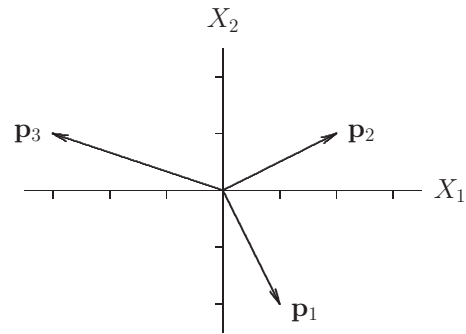
Example: Variable Space

Representing persons in variable space

by points



by vectors



$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \mathbf{p}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{p}_3 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

Person Space

For each test we consider the resulting triples of values

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$$

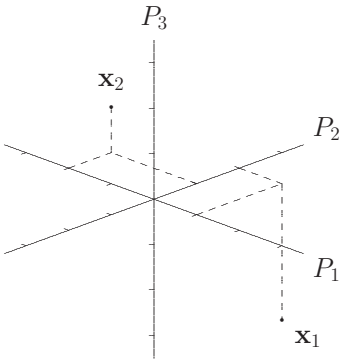
In general, this defines a *person space*

- The axes of the coordinate system correspond to the considered persons
  - In the example results a three-dimensional space with the axes corresponding to the persons P_1, P_2 and P_3
- In this space, each variable (e. g., each diagnostic test) is represented by a point (i. e., by a vector)

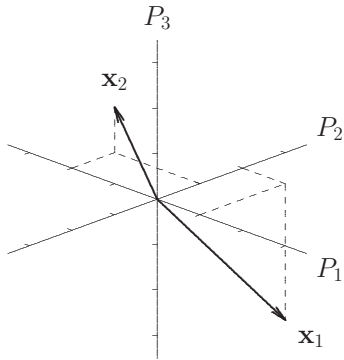
Example: Person Space

Representing variables in person space

by points



by vectors



$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$$

$$\mathbf{x}_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$$

Vectors

Every n-tuple of real numbers $x_1, \ldots, x_n \in \mathbb{R}$ is called a (real) *vector* (n-dimensional vector, n-vector)

- It is called a column vector if it is written as a column of entries

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

- It is called a row vector if it is written as a row of entries

$$(x_1, x_2, \ldots, x_n) \quad \text{or} \quad \begin{pmatrix} x_1 & x_2 & \ldots & x_n \end{pmatrix}$$

Vector Space

- The set of all (real) n-vectors is denoted by $\mathbb{R}^n$
- On the set $\mathbb{R}^n$ we will define operations for n-vectors
  - Addition of vectors
  - Scalar multiplication
- The operations of addition and scalar multiplication satisfy certain properties and impose a structure on $\mathbb{R}^n$
- This structure is called a *(real) vector space*

Vector Addition

For two n-vectors $\mathbf{x}$ and $\mathbf{y}$, their sum $\mathbf{x} + \mathbf{y}$ is defined element-wise by

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

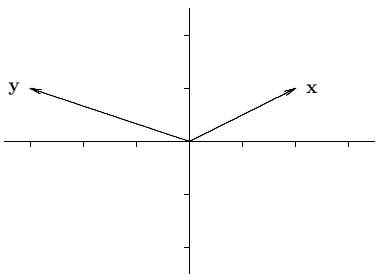
- Geometrically, vector addition corresponds to concatenating or apposing the vectors

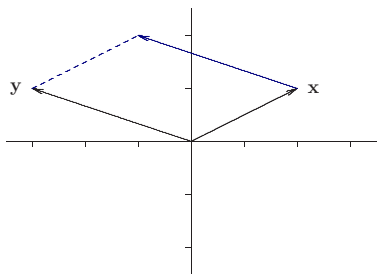
Example: Vector Addition

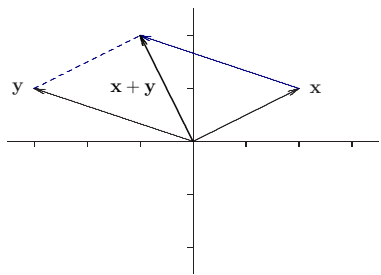
For the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ we obtain the sum

$$\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} 2 + (-3) \\ 1 + 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$







Vector Addition

Vector addition on $\mathbb{R}^n$ defines a *group*

Definition

A *group* is a relational system $\langle G, \oplus \rangle$ consisting of a set G and a binary operation on G, for which the following properties hold

- Associativity: For all $a, b, c \in G$ we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
- There is a neutral element $e \in G$, such that $a \oplus e = e \oplus a = a$ for all $a \in G$
- For each $a \in G$ there exists an inverse element $-a \in G$ with $a \oplus (-a) = (-a) \oplus a = e$

Vector Addition: Group Properties

- Associativity

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$$

for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ is an immediate consequence of associativity of the addition of real numbers holding for each component

- The neutral element is given by the zero vector

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

because obviously $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$

Vector Addition: Group Properties

- For the vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

the inverse element is given by

$$-\mathbf{x} = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}$$

and we have $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$

Scalar Multiplication

For a given n-vector $\mathbf{x} \in \mathbb{R}^n$ and a real number $r \in \mathbb{R}$, the scalar multiplication $r \cdot \mathbf{x}$ is also defined element-wise by

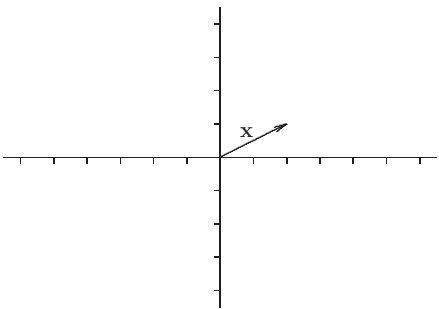
$$r \cdot \mathbf{x} = r \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} r \cdot x_1 \\ r \cdot x_2 \\ \vdots \\ r \cdot x_n \end{pmatrix}$$

- Geometrically, scalar multiplication corresponds to shrinking or stretching the vector

Example: Scalar Multiplication

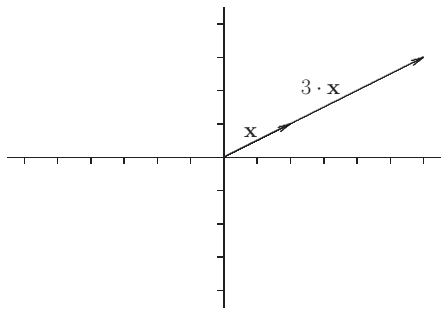
For the vector $\mathbf{x} \in \mathbb{R}^2$

$$\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$



scalar multiplication with $r = 3$ results in

$$3 \cdot \mathbf{x} = \begin{pmatrix} 3 \cdot 2 \\ 3 \cdot 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$



Vector Space

Definition

A real *vector space* is a relational system $\langle V, \oplus, \cdot \rangle$, with $\langle V, \oplus \rangle$ a group satisfying

- Commutativity: For all $\mathbf{a}, \mathbf{b} \in V$ we have $\mathbf{a} \oplus \mathbf{b} = \mathbf{b} \oplus \mathbf{a}$

and $\cdot : \mathbb{R} \times V \to V$ a scalar multiplication such that

- $(r \cdot s) \cdot \mathbf{a} = r \cdot (s \cdot \mathbf{a})$
- $(r + s) \cdot \mathbf{a} = r \cdot \mathbf{a} \oplus s \cdot \mathbf{a}$
- $r \cdot (\mathbf{a} \oplus \mathbf{b}) = r \cdot \mathbf{a} \oplus r \cdot \mathbf{b}$

for all $\mathbf{a}, \mathbf{b} \in V$ and $r, s \in \mathbb{R}$.

- Vector addition and scalar multiplication impose on $\mathbb{R}^n$ the structure of a real vector space

Linear Combinations

For a system of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$ and coefficients $a_1, a_2, \ldots, a_m \in \mathbb{R}$ consider the *linear combination*

$$\sum_{j=1}^{m} a_j \cdot \mathbf{x}_j = a_1 \cdot \mathbf{x}_1 + a_2 \cdot \mathbf{x}_2 + \ldots + a_m \cdot \mathbf{x}_m$$

Linear combinations

- are weighted sums of given vectors
- yield another vector as a result
- are fundamental for introducing the notions of linear independence and dimension of a vector space

Example: Linear Combinations

For the initially considered vectors

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \mathbf{p}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{p}_3 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

and coefficients $a_1 = 2$, $a_2 = 4$ and $a_3 = 0$, we obtain the linear combination

$$\mathbf{p}_4 = \sum_{j=1}^{3} a_j \cdot \mathbf{p}_j = 2 \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} + 4 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}$$

Linear Independence

A system of vectors

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in I\!R^n$$

- is called *linearly independent* if none of the vectors can be generated as a linear combination of the remaining vectors
- has *rank* k whenever k is the maximum number of linearly independent vectors of the given system
- For determining the rank
  - one eliminates vectors, which can be expressed as linear combination of the remaining vectors
  - and repeats this process until all remaining vectors are linearly independent
- The rank of the system is the resulting number $k \leq m$ of linearly independent vectors

Example: Linear Independence

As we can express vector $\mathbf{p}_4$ as a linear combination of the vectors $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$, the system $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$ of vectors

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \mathbf{p}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{p}_3 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \quad \mathbf{p}_4 = \begin{pmatrix} 10 \\ 0 \end{pmatrix}$$

is *not* linearly independent (sometimes, we will also say *linearly dependent*)

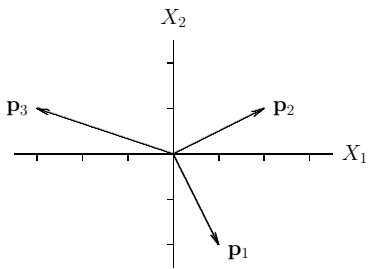
It is easily seen that even the vectors $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ are linearly dependent, but that any two of them form a system of linearly independent vectors

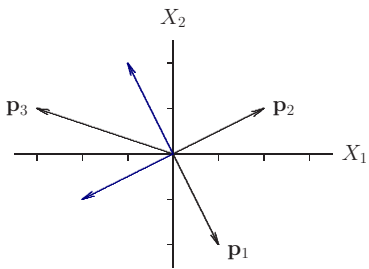
Linear Independence

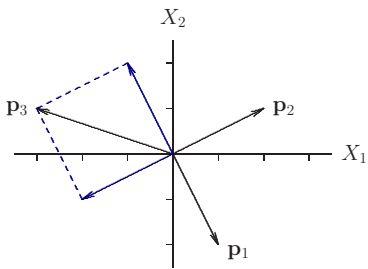
For the vectors

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \mathbf{p}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{p}_3 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

for example, we have $\mathbf{p}_3 = (-1) \cdot \mathbf{p}_1 + (-1) \cdot \mathbf{p}_2$







Linear Independence

Definition

The vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$ are called *linearly independent* if

$$a_1\mathbf{x}_1 + \ldots + a_m\mathbf{x}_m = \mathbf{0}$$

implies that $a_1, \ldots, a_m = 0$ ($\mathbf{0}$ denotes the zero vector)

- Any n linearly independent vectors of $\mathbb{R}^n$ form a *basis* of $\mathbb{R}^n$ (a minimal generating system)
- If $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^n$ form a basis of $\mathbb{R}^n$, then any vector $\mathbf{x} \in \mathbb{R}^n$ can be expressed as a linear combination of these vectors in a unique way
  - For example, $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ form a basis of $\mathbb{R}^2$
- The number n is called the *dimension* of $\mathbb{R}^n$

Vectors in R

Generating vectors

```
x <- c(2, 1)          # R thinks of them as column vectors
y <- c(-3, 1)
```

Vector addition

```
x + y
```

Scalar multiplication

```
3 * x
```

Vectors in R

Linear combinations of vectors

```
p1 <- c(1, -2)
p2 <- c(2, 1)
p3 <- c(-3, 1)
p4 <- 2*p1 + 4*p2 + 0*p3
```

Rank of a system of vectors

```
vs <- cbind(p1, p2, p3, p4)
qr(vs)$rank  # 2
```

- There are different methods for determining the rank of a system of vectors
- The `qr` function computes the QR decomposition (see help page `?qr`)

Matrix

Definition

A rectangular array **A** of $n \cdot m$ real numbers $a_{11}, a_{12}, \ldots, a_{nm}$ with n rows and m columns of the form

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}$$

is called a (real) $(n \times m)$-matrix

Matrix

More efficiently, an $(n \times m)$-matrix

- is denoted as
$$\mathbf{A} = (a_{ij}) \quad i \in \{1, 2, \ldots, n\},\ j \in \{1, 2, \ldots, m\}$$
- and, if it is clear from the context which values the indices i and j can assume, simply by
$$\mathbf{A} = (a_{ij})$$

In the double subscripts of a matrix the row index precedes the column index

Matrix

Notation

- We will denote matrices are by capital letters in boldface (e. g., $\mathbf{A}, \mathbf{X}, \mathbf{S}, \mathbf{\Sigma}, \ldots$) and their entries by lower case letters in normal font with double subscripts
- Vectors are denoted by lower case letters in boldface (e. g., $\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}, \ldots$) and their components by lower case letters in normal font with a single subscript
- By convention, $\mathbf{x}$ denotes a *column vector*

Example

- The matrix representation of the initial example of values of three persons P_1, P_2 und P_3 in two variables X_1 and X_2 leads to the (3×2)-matrix

$$\begin{pmatrix} 1 & -2 \\ 2 & 1 \\ -3 & 1 \end{pmatrix}_{3\times 2}$$

Vectors and Matrices

An $(n \times m)$-matrix can be conceived as being composed of

- m column vectors (n-vectors), or
- n row vectors (m-vectors)

Conversely, any vector can be conceived as a special matrix

- An n-dimensional column vector can be considered as $(n \times 1)$-matrix
- An m-dimensional row vector can be considered as $(1 \times m)$-matrix

Matrices as Linear Mappings

- A matrix can be conceived as a collection of vectors
- A matrix can also be interpreted as a linear mapping (or linear transformation) of vectors
  - An $(n \times m)$-matrix defines a mapping, which assigns to every m-dimensional column vector an n-dimensional column vector
  - This interpretation is based on introducing operations for matrices
    - Addition
    - Scalar multiplication
    - Transposition
    - Matrix multiplication
- Matrices represent structure-preserving mappings between vector spaces (vector space homomorphisms)

Matrix Addition

Given two $(n \times m)$-matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$, the sum $\mathbf{A} + \mathbf{B}$ is defined by

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{pmatrix}$$

Addition is defined for matrices with equal numbers of rows and columns only, and is done element-wise (as with vector addition)

Example: Matrix Addition

Given the two (2×3)-matrices $\mathbf{A}$ and $\mathbf{B}$

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 3 & -4 \\ 1 & 0 & 5 \end{pmatrix}$$

The sum $\mathbf{A} + \mathbf{B}$ is obtained by

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 4+0 & (-2)+3 & 3+(-4) \\ 2+1 & 1+0 & 0+5 \end{pmatrix} = \begin{pmatrix} 4 & 1 & -1 \\ 3 & 1 & 5 \end{pmatrix}$$

Scalar Multiplication

Scalar multiplication with matrices

If $\mathbf{A} = (a_{ij})$ is an $(n \times m)$-matrix and $r \in \mathbb{R}$ a real number, then the scalar multiplication is given by

$$r \cdot \mathbf{A} = \begin{pmatrix} r \cdot a_{11} & r \cdot a_{12} & \dots & r \cdot a_{1m} \\ r \cdot a_{21} & r \cdot a_{22} & \dots & r \cdot a_{2m} \\ \vdots & \vdots & & \vdots \\ r \cdot a_{n1} & r \cdot a_{n2} & \dots & r \cdot a_{nm} \end{pmatrix}$$

Example: Scalar Multiplication

Given the (2×3)-matrix

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix}$$

for the real number $r = 3$ we obtain

$$3 \cdot \mathbf{A} = \begin{pmatrix} 3 \cdot 4 & 3 \cdot (-2) & 3 \cdot 3 \\ 3 \cdot 2 & 3 \cdot 1 & 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 12 & -6 & 9 \\ 6 & 3 & 0 \end{pmatrix}$$

Transposition

Let $\mathbf{A} = (a_{ij})$ be an $(n \times m)$-matrix, then the $(m \times n)$-matrix $\mathbf{A}'$, which results from interchanging rows and columns of $\mathbf{A}$, is called the *transpose* of $\mathbf{A}$

Thus, we have

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}, \quad \mathbf{A}' = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix}$$

Example: Transposition

Given the (2×3)-matrix

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix}$$

The transpose of $\mathbf{A}$ then is the (3×2)-matrix

$$\mathbf{A}' = \begin{pmatrix} 4 & 2 \\ -2 & 1 \\ 3 & 0 \end{pmatrix}$$

Transposition

Symmetric matrices

- An $(n \times n)$-matrix $\mathbf{A}$ is called *symmetric*, whenever

$$\mathbf{A}' = \mathbf{A}$$

- Example

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ -2 & 1 & 0 \\ 3 & 0 & 2 \end{pmatrix}_{3\times 3}$$

- The values symmetrical to the *diagonal* (consisting of elements $a_{11}, a_{22}, \ldots, a_{nn}$) are identical
  - In multivariate statistics, symmetric matrices result, for example, when considering covariances, or correlations, but also for distances

Transposition

Transposition of vectors

Because we can consider an n-dimensional column vector $\mathbf{x}$ as an $(n \times 1)$-matrix, we can also apply transposition to it

The result of transposing the n-dimensional column vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{n \times 1}$$

then is the n-dimensional row vector

$$\mathbf{x}' = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}_{1 \times n}$$

Calculation Rules for Matrices

Concerning the introduced operations for $(n \times m)$-matrices $\mathbf{A}$ and $\mathbf{B}$ and arbitrary real numbers $r, s \in \mathbb{R}$, the following calculation rules hold

- $1 \cdot \mathbf{A} = \mathbf{A}$
- $(r \cdot s) \cdot \mathbf{A} = r \cdot (s \cdot \mathbf{A})$
- $(r + s) \cdot \mathbf{A} = r \cdot \mathbf{A} + s \cdot \mathbf{A}$
- $r \cdot (\mathbf{A} + \mathbf{B}) = r \cdot \mathbf{A} + r \cdot \mathbf{B}$
- $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$
- $(r \cdot \mathbf{A})' = r \cdot \mathbf{A}'$
- $\mathbf{A}'' = (\mathbf{A}')' = \mathbf{A}$

Matrix Multiplication

Let $\mathbf{A} = (a_{ik})$ be an $(n \times l)$-matrix and $\mathbf{B} = (b_{kj})$ an $(l \times m)$-matrix

The $(n \times m)$-matrix $\mathbf{C} = (c_{ij})$ is called the *product* of $\mathbf{A}$ and $\mathbf{B}$ if each element c_{ij} is given by

$$c_{ij} = \sum_{k=1}^{l} a_{ik} \cdot b_{kj}$$

In this case the notations

$$\mathbf{C} = \mathbf{AB}$$

or (with explicitly indicating rows and columns)

$$\mathbf{C}_{(n \times m)} = \mathbf{A}_{(n \times l)} \mathbf{B}_{(l \times m)}$$

are used

Matrix Multiplication

Note that the product of two matrices is defined only if the number of columns of the first matrix is the same as the number of rows of the second matrix

Learning rule

- Calculation of the elements c_{ij} in the i-th row and j-th column of the product matrix $\mathbf{C} = \mathbf{AB}$ proceeds by considering the elements a_{ik} and b_{kj} for $k \in \{1, \ldots, l\}$
- We can thus say that the i-th row of matrix $\mathbf{A}$ is "multiplied" with the j-th column of the matrix $\mathbf{B}$

Matrix Multiplication

"Multiplication" of the i-th row of the $(n \times l)$-matrix $\mathbf{A}$ with the j-th column of the $(l \times m)$-matrix $\mathbf{B}$

$$\mathbf{C} = \mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2k} & \dots & a_{2l} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ik} & \dots & a_{il} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} & \dots & a_{nl} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{k1} & b_{k2} & \dots & b_{kj} & \dots & b_{km} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{l1} & b_{l2} & \dots & b_{lj} & \dots & b_{lm} \end{pmatrix}$$

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{il}b_{lj}$$

Example: Matrix Multiplication

Given the (2×3)-matrix

$$\mathbf{A}_{(2\times3)} = \begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix}$$

and the (3×2)-matrix

$$\mathbf{B}_{(3\times2)} = \begin{pmatrix} 1 & 3 \\ 4 & -1 \\ 0 & 2 \end{pmatrix}$$

The product $\mathbf{C} = \mathbf{AB}$ then is a (2×2)-matrix

$$\mathbf{C}_{(2\times2)} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

$$= \begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & -1 \\ 0 & 2 \end{pmatrix}$$

$$\mathbf{A}_{(2\times3)}\mathbf{B}_{(3\times2)} = \mathbf{C}_{(2\times2)}$$

Example: Matrix Multiplication

"Multiplication" of the 1st row of **A** with the 1st column of **B**

$$\begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & -1 \\ 0 & 2 \end{pmatrix}$$

$$\begin{aligned} c_{11} &= a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} \\ &= 4 \cdot 1 + (-2) \cdot 4 + 3 \cdot 0 \\ &= -4 \end{aligned}$$

"Multiplication" of the 1st row of **A** with the 2nd column of **B**

$$\begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & -1 \\ 0 & 2 \end{pmatrix}$$

$$\begin{aligned} c_{12} &= a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + a_{13} \cdot b_{32} \\ &= 4 \cdot 3 + (-2) \cdot (-1) + 3 \cdot 2 \\ &= 20 \end{aligned}$$

Example: Matrix Multiplication

"Multiplication" of the 2nd row of **A** with the 1st column of **B**

$$\begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & -1 \\ 0 & 2 \end{pmatrix}$$

$$\begin{aligned} c_{21} &= a_{21} \cdot b_{11} + a_{22} \cdot b_{21} + a_{23} \cdot b_{31} \\ &= 2 \cdot 1 + 1 \cdot 4 + 0 \cdot 0 \\ &= 6 \end{aligned}$$

"Multiplication" of the 2nd row of **A** with the 2nd column of **B**

$$\begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & -1 \\ 0 & 2 \end{pmatrix}$$

$$\begin{aligned} c_{22} &= a_{21} \cdot b_{12} + a_{22} \cdot b_{22} + a_{23} \cdot b_{32} \\ &= 2 \cdot 3 + 1 \cdot (-1) + 0 \cdot 2 \\ &= 5 \end{aligned}$$

Example: Matrix Multiplication

Thus, for the product $\mathbf{C} = \mathbf{AB}$ we get

$$\mathbf{C} = \begin{pmatrix} 4 & -2 & 3 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} -4 & 20 \\ 6 & 5 \end{pmatrix}$$

Matrix Multiplication

If the product **AB** is defined, then the product **BA** exists only if the number of columns of **B** equals the number of rows of **A**

If **A** is an $(n \times m)$-matrix and **B** an $(m \times n)$-matrix, then both products **AB** and **BA** are defined

- Since **AB** is an $(n \times n)$-matrix and **BA** an $(m \times m)$-matrix, it is obvious that $\mathbf{AB} \neq \mathbf{BA}$ whenever $m \neq n$
- However, even in case of $m = n$, in general $\mathbf{AB} \neq \mathbf{BA}$

Example: Matrix Multiplication

For

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 3 & 5 \\ 5 & 6 \end{pmatrix}$$

we have

$$\mathbf{AB} = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 11 & 16 \\ 23 & 29 \end{pmatrix}$$

and

$$\mathbf{BA} = \begin{pmatrix} 3 & 5 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 11 & 23 \\ 16 & 29 \end{pmatrix}$$

Matrix Multiplication

Calculation rules

Let $\mathbf{A}, \mathbf{A}_0$ be two $(n \times l)$-matrices, $\mathbf{B}, \mathbf{B}_0$ two $(l \times m)$-matrices, and $r \in \mathbb{R}$ a real number

Besides associativity of matrix multiplication the following properties hold

- $\mathbf{A}(\mathbf{B} + \mathbf{B}_0) = \mathbf{AB} + \mathbf{AB}_0$
- $(\mathbf{A} + \mathbf{A}_0)\mathbf{B} = \mathbf{AB} + \mathbf{A}_0\mathbf{B}$
- $\mathbf{A}(r \cdot \mathbf{B}) = (r \cdot \mathbf{A})\mathbf{B} = r \cdot (\mathbf{AB})$
- $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

Square Matrix

An $(n \times n)$-matrix $\mathbf{A} = (a_{ij})$, with equal number of rows and columns, is called a *square matrix*, and the collection of elements a_{ii} mit $i = 1, \ldots, n$ is its *diagonal*

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1i} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2i} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{ii} & \ldots & a_{in} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{ni} & \ldots & a_{nn} \end{pmatrix}$$

Diagonal Matrix

If an $(n \times n)$ square matrix $\mathbf{A} = (a_{ij})$ satisfies $a_{ij} = 0$ for all $i \neq j$, then in is called a *diagonal matrix*

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{ii} & \dots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & a_{nn} \end{pmatrix}$$

Identity Matrix

An $(n \times n)$ diagonal matrix $\mathbf{A} = (a_{ij})$, with $a_{ii} = 1$ for all i, is called *identity matrix* (or $(n \times n)$-*identity matrix*) and denoted by $\mathbf{I}_n$ (or $\mathbf{I}$)

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

Identity Matrix and Inverse

In matrix multiplication, the identity matrix plays the role of the one, i. e., of the neutral element

- For all $(n \times m)$-matrices $\mathbf{A}$

$$\mathbf{I}_n\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_m$$

If for a square matrix $\mathbf{A}$ there is a square matrix $\mathbf{B}$ with

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I},$$

then $\mathbf{A}$ is called *invertible* (or *non-singular*, *regular*)

The matrix $\mathbf{B}$, which is uniquely determined by these equations, is called the *inverse* of $\mathbf{A}$ and is denoted by $\mathbf{A}^{-1}$

Matrices in R

Generating matrices

```
A <- matrix(c(4, -2, 3,  2, 1, 0), 2, 3, byrow = TRUE)  # fill by row
B <- matrix(c(0, 3, -4,  1, 0, 5), 2, 3, byrow = TRUE)
```

Addition of matrices

```
A + B
```

Scalar multiplication

```
3 * A
```

Matrices in R

Transposition

```
t(A)      # transpose of A = A'
t(t(A))   # A'' = A
t(x)      # row vector
```

Matrix multiplication

```
B <- matrix(c(1, 3,  4, -1,  0, 2), 3, 2, byrow = TRUE)
A %*% B     # (2 x 2)-matrix
A * t(B)    # element-wise multiplication: (2 x 3)-matrix
```

Matrix multiplication need not be commutative

```
A <- matrix(c(2, 1,  1, 4), 2, 2, byrow = TRUE)
B <- matrix(c(3, 5,  5, 6), 2, 2, byrow = TRUE)
A %*% B
B %*% A
```

Matrices in R

Diagonal matrix

```
diag(1:5)      # diagonal matrix with elements 1, ..., 5
diag(2.5, 4)   # diagonal matrix with element 2.5
```

Extract the diagonal

```
A <- matrix(1:9, 3, 3)
diag(A)
```

Identity matrix

```
diag(6)        # I_6
```

Characteristic Values of Square Matrices

Two characteristic values of square matrices useful for multivariate statistics

- trace
- determinant

These values remain invariant under frequently occurring transformations

Moreover, the determinant indicates whether a matrix is invertible

Trace

The trace of a square matrix is the sum of its diagonal elements

- For an $(n \times n)$ square matrix $\mathbf{A}$, the trace $\mathrm{tr}(\mathbf{A})$ is defined by

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$$

- Calculation rules
  - $\mathrm{tr}(\mathbf{I_n}) = n$
  - $\mathrm{tr}(\mathbf{A}') = \mathrm{tr}(\mathbf{A})$
  - $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
  - For an $(n \times m)$-matrix $\mathbf{A}$ and an $(m \times n)$-matrix $\mathbf{B}$ we have $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

Example: Trace

- Given the (3×3)-matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 1 \\ 5 & 2 & 3 \end{pmatrix}$$

- The trace of $\mathbf{A}$ then is

$$\mathrm{tr}(\mathbf{A}) = a_{11} + a_{22} + a_{33} = 2 + 4 + 3 = 9$$

Determinant

The determinant of an $(n \times n)$-matrix $\mathbf{A}$ is defined by induction

- It is denoted by $\det(\mathbf{A})$ or by $|\mathbf{A}|$
- For a (1×1)-matrix $\mathbf{A} = (a_{11})$ we have

$$\det(\mathbf{A}) = a_{11}$$

- For a (2×2)-matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

we have

$$\det(\mathbf{A}) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$$

Determinant

Calculation

- In general, and based on these equations, the determinant is calculated according to the following recursive rule
- For an $(n \times n)$-matrix $\mathbf{A}$ let

$$\mathbf{A}_{ij}$$

denote the $((n-1) \times (n-1))$-matrix, which results from deleting the i-th row and the j-th column in $\mathbf{A}$

- Then one defines

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(\mathbf{A}_{ij})$$

(Developing the determinant along the i-th row)

Determinant

Calculation rules (for square matrices)

- $\det(\mathbf{A}') = \det(\mathbf{A})$
- For $(n \times n)$-matrices $\mathbf{A} = (a_{ik})$ and $\mathbf{B} = (b_{kj})$ we have

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$$

- $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1} = 1/\det(\mathbf{A})$
- For a diagonal matrix $\mathbf{D}$ the determinant equals the product of the diagonal elements

$$\det(\mathbf{D}) = \prod_{i=1}^{n} d_{ii}$$

- Thus, we have $\det(\mathbf{I}) = 1$

The determinant of a square matrix is non-zero if and only if the matrix is invertible

Example: Determinant

- The determinant of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 1 \\ 5 & 2 & 3 \end{pmatrix}$$

  will be developed along the 1st row

- This yields

$$\begin{aligned} \det(\mathbf{A}) &= \sum_{j=1}^{n} (-1)^{1+j} \cdot a_{1j} \cdot \det(\mathbf{A}_{1j}) \\ &= (-1)^{1+1} \cdot a_{11} \cdot \det(\mathbf{A}_{11}) + (-1)^{1+2} \cdot a_{12} \cdot \det(\mathbf{A}_{12}) \\ &\quad + (-1)^{1+3} \cdot a_{13} \cdot \det(\mathbf{A}_{13}) \\ &= a_{11} \cdot \det(\mathbf{A}_{11}) - a_{12} \cdot \det(\mathbf{A}_{12}) + a_{13} \cdot \det(\mathbf{A}_{13}) \end{aligned}$$

Example: Determinant

- Where

$$\begin{aligned}
\det(\mathbf{A}_{11}) &= \det\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} = 4 \cdot 3 - 1 \cdot 2 = 10 \\
\det(\mathbf{A}_{12}) &= \det\begin{pmatrix} 1 & 1 \\ 5 & 3 \end{pmatrix} = 1 \cdot 3 - 1 \cdot 5 = -2 \\
\det(\mathbf{A}_{13}) &= \det\begin{pmatrix} 1 & 4 \\ 5 & 2 \end{pmatrix} = 1 \cdot 2 - 4 \cdot 5 = -18
\end{aligned}$$

- Finally, the determinant of $\mathbf{A}$ is given by

$$\begin{aligned}
\det(\mathbf{A}) &= a_{11} \cdot \det(\mathbf{A}_{11}) - a_{12} \cdot \det(\mathbf{A}_{12}) + a_{13} \cdot \det(\mathbf{A}_{13}) \\
&= 2 \cdot 10 - 3 \cdot (-2) + 0 \cdot (-18) \\
&= 26
\end{aligned}$$

Example: Determinant

- Because the determinant $\det(\mathbf{A})$ is non-zero, the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 1 \\ 5 & 2 & 3 \end{pmatrix}$$

  is invertible
- Later we show that

$$\mathbf{A}^{-1} = \frac{1}{26} \begin{pmatrix} 10 & -9 & 3 \\ 2 & 6 & -2 \\ -18 & 11 & 5 \end{pmatrix}$$

- Check that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$

Determinant

Rule of Sarrus

- For a (3×3)-matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

  the calculation of the determinant can be simplified

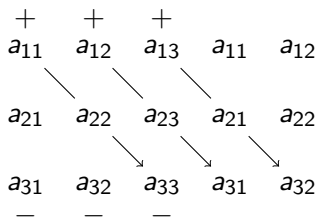
- Following the *rule of Sarrus* one calculates

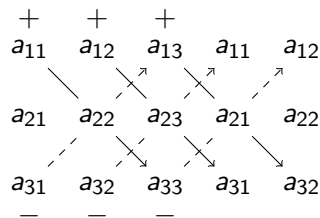
$$\begin{aligned} \det(\mathbf{A}) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12} \end{aligned}$$

Determinant

Rule of Sarrus

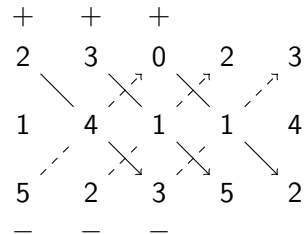
- Procedure
  - Write out the first two columns of the matrix to the right of the third column
  - Then add the products of the diagonals going from top to bottom (solid) and subtract the products of the diagonals going from bottom to top (dashed)





Example: Determinant

- The rule of Sarrus provides



$$\det\begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 1 \\ 5 & 2 & 3 \end{pmatrix} = 2\cdot 4\cdot 3 + 3\cdot 1\cdot 5 + 0\cdot 1\cdot 2$$
$$- \; 5\cdot 4\cdot 0 - 2\cdot 1\cdot 2 - 3\cdot 1\cdot 3 = 26$$

Matrix Inversion

Inverse of a matrix

- A matrix $\mathbf{A}$ is called *invertible* if there is a matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{I} \quad \text{and} \quad \mathbf{BA} = \mathbf{I}$$

- The matrix $\mathbf{B}$, which is uniquely determined by these equations, is called the *inverse* of $\mathbf{A}$
- The inverse of $\mathbf{A}$ is denoted by $\mathbf{A}^{-1}$ and we have

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$

- Matrices that are not invertible are also called *singular*

Matrix Inversion

Rank of a matrix

- In order to determine whether a square matrix $\mathbf{A}$ has an inverse, we refer to the notion of the rank of a system of vectors
- In general, for any $(n \times m)$-matrix $\mathbf{A}$ the row rank is defined as the rank of the system of row vectors constituting the matrix (the column rank is defined analogously)
- It can be shown that the row rank always equals the column rank, so we can consider the *rank* $rk(\mathbf{A})$ of the matrix $\mathbf{A}$
- We have $rk(\mathbf{A}) \leq \min\{n, m\}$
- Calculation rules
  - $rk(\mathbf{A}) = rk(\mathbf{A}')$
  - $rk(\mathbf{A}'\mathbf{A}) = rk(\mathbf{A}) = rk(\mathbf{A}\mathbf{A}')$

Elementary Operations

Rank of a matrix

- The rank of a matrix can be calculated by *elementary operations*
- Elementary row operations of a matrix **A** include
  1. Row switching: interchanging rows
  2. Row multiplication: multiplying a row with a scalar $a \neq 0$
  3. Row addition: add a multiple of another row to a particular row
- Goal: By repeated application of elementary row operations transform the matrix **A** into *row echelon form*
- The row echelon form allows for reading off the rank of **A**

Row Echelon Form

A matrix $\tilde{\mathbf{A}}$ in row echelon form looks like this

$$\tilde{\mathbf{A}} = \begin{pmatrix} \bullet & * & * & * & * & * & \cdots & * & \cdots & * \\ 0 & 0 & \bullet & * & * & * & \cdots & * & \cdots & * \\ \vdots & & 0 & \bullet & * & * & & \vdots & & \vdots \\ & & \vdots & 0 & 0 & \bullet & & & & \\ & & & & & & & & & \\ 0 & & 0 & \cdots & & 0 & \bullet & * & \cdots & * \\ 0 & & 0 & & & & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & & & & \vdots & \vdots & & \vdots \\ 0 & & 0 & & & & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$\bullet$-entries: numbers $\neq 0$, $*$-entries: arbitrary numbers resulting from row operations, all other entries: 0

The number of rows of $\tilde{\mathbf{A}}$ with a $\bullet$-entry provides the rank of $\mathbf{A}$

Example: Rank

Determining the rank of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 2 \\ 2 & 12 & 7 \\ 1 & 10 & 6 \end{pmatrix}$$

Row echelon form of $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 2 \\ 2 & 12 & 7 \\ 1 & 10 & 6 \end{pmatrix} \xrightarrow[\text{III}-\frac{1}{2}\text{I}]{\text{II}-\text{I}} \begin{pmatrix} 2 & 4 & 2 \\ 0 & 8 & 5 \\ 0 & 8 & 5 \end{pmatrix} \xrightarrow{\text{III}-\text{II}} \begin{pmatrix} 2 & 4 & 2 \\ 0 & 8 & 5 \\ 0 & 0 & 0 \end{pmatrix} = \tilde{\mathbf{A}}$$

$\mathbf{A}$ has rank 2

Matrix Inversion: Equivalences

For an $(n \times n)$ square matrix $\mathbf{A}$ the following statements are equivalent (each of the statements is true if and only if the others are true)

- $\mathbf{A}$ is invertible
- The rank of $\mathbf{A}$ equals n (we say $\mathbf{A}$ has full rank)
- $\det(\mathbf{A}) \neq 0$
- The row vectors and the column vectors of $\mathbf{A}$, respectively, are linearly independent

Matrix Inversion

Calculation rules

- If the matrices $\mathbf{A}$ and $\mathbf{B}$ are invertible, then this also holds for $\mathbf{AB}$ and we have

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

- If $\mathbf{A}$ is invertible, so is its transpose $\mathbf{A}'$, and we get

$$(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$$

Matrix Inversion

Calculating the inverse

- Various algorithms have been developed for determining the inverse of an invertible square matrix
- These include the Gauß-Jordan algorithm (or Gaussian elimination) and the method of cofactors
- The calculations can become tedious already for matrices of moderate size, so that applying computer programs is recommended (e. g., R or SageMath)
- The calculations, however, are fairly simple for special matrices

Matrix Inversion

Diagonal matrix

- A diagonal matrix $\mathbf{A}$ is invertible if and only if all diagonal elements are non-zero
- The inverse $\mathbf{A}^{-1}$ then is the diagonal matrix with reciprocals of the diagonal elements of $\mathbf{A}$
- Example: The inverse of the matrix

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

is

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

Matrix Inversion

(2×2)-matrix

- For an invertible (2×2)-matrix $\mathbf{A}$ we have

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

- Example

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{-1} = \frac{1}{1 \cdot 4 - 3 \cdot 2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1.5 & -0.5 \end{pmatrix}$$

- Check

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 1.5 & -0.5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Matrix Inversion

General procedure

- The inverse of a matrix $\mathbf{A}$ can be calculated via its cofactors (minors)
- Let $a_{ij}^{(-1)}$ denote the entry in $\mathbf{A}^{-1}$ in the i-th row and the j-th column, and let

$$A_{ji} = (-1)^{j+i} \cdot \det(\mathbf{A}_{ji})$$

  be the *cofactor* of a_{ji}
- Here, $\mathbf{A}_{ji}$ is the matrix that results from deleting the j-th row and the i-th column in $\mathbf{A}$ (cf. developing the determinant)
- Then the inverse $\mathbf{A}^{-1}$ is developed element-wise by

$$a_{ij}^{(-1)} = \frac{A_{ji}}{\det(\mathbf{A})}$$

Example: Cofactor Method

(3×3)-matrix

- Developing the inverse $\mathbf{A}^{-1}$ of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 1 \\ 5 & 2 & 3 \end{pmatrix}$$

by applying the cofactor method via

$$a_{ij}^{(-1)} = \frac{A_{ji}}{\det(\mathbf{A})}$$

- We first calculate the determinant

$$\det(\mathbf{A}) = 2 \cdot 4 \cdot 3 + 3 \cdot 1 \cdot 5 + 0 - 0 - 2 \cdot 1 \cdot 2 - 3 \cdot 1 \cdot 3 = 26$$

Example: Cofactor Method

(3×3)-matrix

- Then the cofactors $A_{ji} = (-1)^{j+i} \cdot \det(\mathbf{A}_{ji})$ are determined element-wise by

$$
\begin{aligned}
a_{11}^{(-1)}: &\qquad A_{11} = (-1)^{1+1} \cdot \det(\mathbf{A}_{11}) = +(4 \cdot 3 - 2 \cdot 1) = 10 \\
a_{12}^{(-1)}: &\qquad A_{21} = (-1)^{2+1} \cdot \det(\mathbf{A}_{21}) = -9 \\
a_{13}^{(-1)}: &\qquad A_{31} = (-1)^{3+1} \cdot \det(\mathbf{A}_{31}) = +3 \\
\\
a_{21}^{(-1)}: &\qquad A_{12} = (-1)^{1+2} \cdot \det(\mathbf{A}_{12}) = -(-2) \\
a_{22}^{(-1)}: &\qquad A_{22} = (-1)^{2+2} \cdot \det(\mathbf{A}_{22}) = +6 \\
a_{23}^{(-1)}: &\qquad A_{32} = (-1)^{3+2} \cdot \det(\mathbf{A}_{32}) = -2
\end{aligned}
$$

Example: Cofactor Method

(3×3)-matrix

- Cofactors (continued)

$$
\begin{aligned}
a_{31}^{(-1)}: \quad & A_{13} = (-1)^{1+3} \cdot \det(\mathbf{A}_{13}) = +(-18) \\
a_{32}^{(-1)}: \quad & A_{23} = (-1)^{2+3} \cdot \det(\mathbf{A}_{23}) = -(-11) \\
a_{33}^{(-1)}: \quad & A_{33} = (-1)^{3+3} \cdot \det(\mathbf{A}_{33}) = +5
\end{aligned}
$$

- This provides

$$
\mathbf{A}^{-1} = \frac{1}{26} \begin{pmatrix} 10 & -9 & 3 \\ 2 & 6 & -2 \\ -18 & 11 & 5 \end{pmatrix} \approx \begin{pmatrix} 0.385 & -0.346 & 0.115 \\ 0.077 & 0.231 & -0.077 \\ -0.692 & 0.423 & 0.192 \end{pmatrix}
$$

Scalar Product

Definition

- For given n-dimensional column vectors $\mathbf{x}$ and $\mathbf{y}$ the *scalar product* (or *dot product*) $\langle \mathbf{x}, \mathbf{y} \rangle$ can be defined as a special case of matrix multiplication by

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle &= \mathbf{x}'\mathbf{y} \\ &= \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \sum_{i=1}^{n} x_i \cdot y_i \end{aligned}$$

Norm

Given orthogonal coordinate axes with equal units there is the following interpretation

- By repeated application of the Pythagorean theorem one realizes that

$$\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} x_i^2 = |\mathbf{x}|^2$$

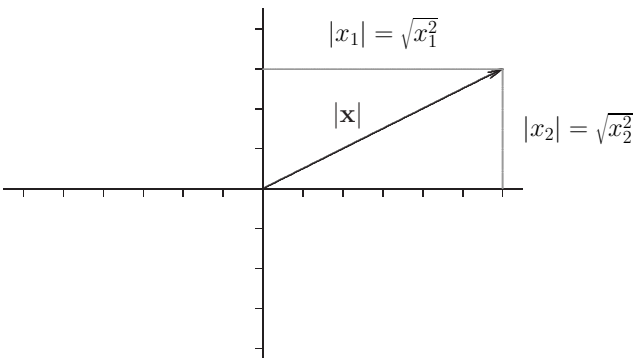
  is the squared length of the vector $\mathbf{x}$
- The *length* (or *norm*) of the vector $\mathbf{x}$ is then given by

$$|\mathbf{x}| = \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

  (*Euclidean norm*)

Norm

Graphical illustration of the length of a vector in $\mathbb{R}^2$



The length of $\mathbf{x} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$ is

$$\left| \begin{pmatrix} 6 \\ 3 \end{pmatrix} \right| = \sqrt{\begin{pmatrix} 6 & 3 \end{pmatrix} \begin{pmatrix} 6 \\ 3 \end{pmatrix}}$$

$$= \sqrt{6^2 + 3^2} = \sqrt{45}$$

$$\approx 6.71$$

Norm

Unit vector

- In a vector space one often considers vectors of length 1 (e. g., when specifying a basis)
- A vector $\mathbf{e}$ with $|\mathbf{e}| = 1$ is called *unit vector* (or *normalized vector*)
- For a vector $\mathbf{x}$ the corresponding unit vector is given by

$$\frac{\mathbf{x}}{|\mathbf{x}|}$$

- The unit vector has the same spatial direction as the vector $\mathbf{x}$

Metric

Based on a scalar product (and thus a norm) one can introduce a metric (or distance function)

- A *metric* on $\mathbb{R}^n$ is defined by

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and satisfies

  - $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$ (minimality)
  - $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (symmetry)
  - $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ for all $\mathbf{z} \in \mathbb{R}^n$ (triangle inequality)

- These properties imply $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y}$
- In the current case, we get the *Euclidean Metric*

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Orthogonality

Given orthogonal coordinate axes with equal units, the scalar product of two vector has a geometrical interpretation

- The number

$$\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{|\mathbf{x}| \cdot |\mathbf{y}|}$$

  specifies the cosine of the angle enclosed by the vectors $\mathbf{x}$ and $\mathbf{y}$
- Therefore, the vectors $\mathbf{x}$ und $\mathbf{y}$ are orthogonal if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

Linear Mappings

Matrices were already characterized as linear mappings (or linear transformations)

- An $(n \times m)$-matrix $\mathbf{A}$ represents a linear mapping of the m-dimensional vector space $\mathbb{R}^m$ into the n-dimensional vector space $\mathbb{R}^n$
  - To each m-vector $\mathbf{x}$ the n-vector $\mathbf{Ax}$ is assigned

$$\mathbf{x} \mapsto \mathbf{Ax}$$

  - Linearity means that for all m-vectors $\mathbf{x}, \mathbf{y}$ and real numbers r

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay}$$
$$\mathbf{A}(r \cdot \mathbf{x}) = r \cdot (\mathbf{Ax})$$

Linear Mappings

Interpretation

- A characteristic feature of linear mappings is that they map lines onto lines
  - Examples of linear mappings are reflections, rotations, homogeneous dilations ...
- Linear mappings play an important role when transforming variables, i. e., the transition between different coordinate systems (e. g., in principle components and factor analysis)
- The matrix product $\mathbf{AB}$ corresponds to the composition of the linear mappings associated to the matrices $\mathbf{B}$ and $\mathbf{A}$

Systems of Linear Equations

The multiplication of a matrix with a vector has the following basic interpretation

- If $\mathbf{x}$ is a m-vector, $\mathbf{y}$ a n-vector and $\mathbf{A}$ a $(n \times m)$-matrix then

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

characterizes the system of linear equations

$$\begin{aligned} y_1 &= a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1m}\,x_m \\ y_2 &= a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2m}\,x_m \\ &\vdots \\ y_n &= a_{n1}\,x_1 + a_{n2}\,x_2 + \ldots + a_{nm}\,x_m \end{aligned}$$

Solving Systems of Linear Equations

Often one encounters the special situation that the n-vector $\mathbf{y}$ and the $(n \times n)$ square matrix $\mathbf{A}$ are given

- We want to find the n-vectors $\mathbf{x}$, for which the equation

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

  holds

- It depends on the matrix $\mathbf{A}$, how many solutions there are and whether solutions exist at all (how many different vectors $\mathbf{x}$ exist that satisfy the equation)
  - In principle there are n equations in n unknowns $x_1, \ldots, x_n$, which are to be determined
  - The following provides a general criterion for simultaneously solving all n equations

Solving Systems of Linear Equations

- If the matrix $\mathbf{A}$ is invertible then the system

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

of n equations in n unknowns has a unique solution (there is exactly one vector $\mathbf{x}$ for which the equation holds)

- Multiplying the given equation from the left with the inverse $\mathbf{A}^{-1}$ provides

$$\mathbf{A}^{-1}\mathbf{y} = \mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x}$$

- The unique solution $\mathbf{x}$ of the system of equations then is given by

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

Solving Systems of Linear Equations

- If the matrix $\mathbf{A}$ is *not invertible*, the equation $\mathbf{y} = \mathbf{A}\mathbf{x}$ has either no solution at all or there exist more than one solution
- In this case $\mathbf{A}$ maps all vectors $\mathbf{x}$ into a subspace of $\mathbb{R}^n$ (a subspace with dimension $< n$)
  - If $\mathbf{y}$ is *not* an element of this subspace then the system of equations has no solution
  - If $\mathbf{y}$ is an element of this subspace then several (in fact, infinitely many) solutions to the system of equations exist (less equations than unknowns)
- Formally, there exist solutions if the rank of the *coefficient matrix* $\mathbf{A}$ equals the rank of the *extended coefficient matrix* $(\mathbf{A}|\mathbf{y})$
- The solution is unique if these matrices have full rank n

Characteristic Values of Square Matrices in R

Trace

```
A <- matrix(c(2, 3, 0,
              1, 4, 1,
              5, 2, 3), 3, 3, byrow = TRUE)
sum(diag(A))
```

Determinant

```
det(A)
```

Matrix Inversion in R

Inverse of a diagonal matrix

```
A <- diag(c(2, 4, 3))
solve(A)
solve(A) |> MASS::fractions()
```

Inverse of a general square matrix

```
A <- matrix(c(2, 3, 0,                          # (3 x 3)-matrix
              1, 4, 1,
              5, 2, 3), 3, 3, byrow = TRUE)
solve(A)
solve(A) |> MASS::fractions()
A %*% solve(A)                                  # check AA^{-1} = I_3
solve(A) %*% A                                  # check A^{-1}A = I_3
```

Vector Products in R

Scalar product $\mathbf{x}'\mathbf{x}$

```
x <- c(6, 3)
t(x) %*% x
sqrt(t(x) %*% x)  # norm
```

Dyadic (outer) product $\mathbf{x}\mathbf{x}'$

```
x %*% t(x)        # (2 x 2)-matrix
```

Solving Systems of Equations with R

Solving systems of equations $\mathbf{Ax} = \mathbf{y}$

```
A <- matrix(c(2, 1, 1, -2), 2, 2, byrow = TRUE)
y <- c(-2, 4)
solve(A, y)            # x = A^{-1}y

A <- matrix(c(1, -2, -2, 4), 2, 2, byrow = TRUE)
y <- c(2, 1)
solve(A, y)            # A is singular
```

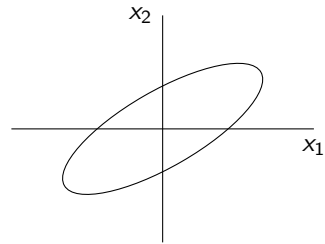
Are there no or several solutions?

```
qr(A)$rank
qr(cbind(A, y))$rank   # rk(A) < rk(A|y), no solution
```

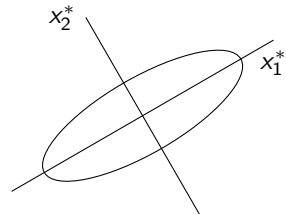
Eigenvalues and Eigenvectors

Eigenvalue problem: geometric interpretation

- Let the data points in coordinates (x_1, x_2) be correlated



- Wanted: new coordinates (x_1^*, x_2^*) such that the data points are uncorrelated



Eigenvalues and Eigenvectors

Definition

- In general, one considers n-vectors $\mathbf{x}$ for which multiplication with a symmetric $(n \times n)$-matrix $\mathbf{A}$ has the same effect as multiplication with a scalar λ (real or complex), such that
$$\mathbf{A}\mathbf{x} = \lambda \cdot \mathbf{x}$$
- If there is such a value λ for any vector $\mathbf{x}$ then it is called an *eigenvalue* of the matrix $\mathbf{A}$
- The vector $\mathbf{x}$ then is called an *eigenvector* corresponding to the eigenvalue λ

Eigenvalues and Eigenvectors

Calculation

- We are looking for vectors $\mathbf{x}$, such that for a given symmetric $(n \times n)$-matrix $\mathbf{A}$ there is a scalar λ with

$$\mathbf{A}\mathbf{x} = \lambda \cdot \mathbf{x}$$

- This means that

$$\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$$

and, factoring out $\mathbf{x}$,

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

Eigenvalues and Eigenvectors

Calculation

- The equation

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

  represents a system of linear equations of the form $\mathbf{Bx} = \mathbf{y}$, which has a unique solution if and only if $\mathbf{B} = \mathbf{A} - \lambda \mathbf{I}_n$ is invertible

- If the matrix $\mathbf{A} - \lambda \mathbf{I}_n$ is invertible then

$$\begin{aligned} \mathbf{x} &= (\mathbf{A} - \lambda \mathbf{I}_n)^{-1}\mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

  and only the zero-vector $\mathbf{x} = \mathbf{0}$ satisfies the equation

Eigenvalues and Eigenvectors

Calculation

- For non-zero vectors $\mathbf{x} \neq \mathbf{0}$ satisfying the equation

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

  to exist, the matrix $\mathbf{A} - \lambda \mathbf{I}_n$ needs to be *singular*
- This means that

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

- This equation is called the *characteristic equation* of the matrix $\mathbf{A}$

Eigenvalues and Eigenvectors

Calculation

- In case of $n = 2$, one considers

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}_2) &= \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12} a_{21} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11} a_{22} - a_{12} a_{21} \end{aligned}$$

- The *eigenvalues* λ then are the roots (zeros) of the resulting polynomial of degree 2

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11} a_{22} - a_{12} a_{21}$$

(in general a polynomial of degree n results, the *characteristic polynomial*)

Example: Eigenvalues and Eigenvectors

- The characteristic polynomial of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$

reads

$$\det \begin{pmatrix} 1-\lambda & 0.8 \\ 0.8 & 1-\lambda \end{pmatrix} = \lambda^2 - 2\lambda + 0.36$$

- Because a polynomial of degree n has n (not necessarily distinct, and maybe complex) roots, the equation

$$\lambda^2 - 2\lambda + 0.36 = 0$$

results in the two eigenvalues

$$\lambda_1 = 1.8 \quad \text{and} \quad \lambda_2 = 0.2$$

Eigenvectors

Calculation

- Given the eigenvalue λ of the matrix $\mathbf{A}$, the vectors $\mathbf{x}$ satisfying

$$\mathbf{A}\mathbf{x} = \lambda \cdot \mathbf{x}$$

  are the *eigenvectors* corresponding to the eigenvalue λ
- Together with the zero vector $\mathbf{0}$, the eigenvectors corresponding to a particular eigenvalue form a linear subspace of the vector space, the *eigenspace* (dimension = multiplicity of the root)
- Eigenvectors corresponding to distinct eigenvalues are orthogonal

Eigenvectors

Calculation

- Eigenvectors are obtained after plugging in the eigenvalue by computing the *adjunct matrix* (also *adjugate* or *classical adjoint*)

$$adj(\mathbf{A} - \lambda\mathbf{I})$$

  which is related to the cofactors of $\mathbf{A} - \lambda\mathbf{I}$ (cf. matrix inversion: $\mathbf{B}^{-1} = adj(\mathbf{B})/\det\mathbf{B}$)
- Eigenvectors are unique up to multiplication with a constant, which means that all proportional columns of the adjunct matrix are eigenvectors to a given eigenvalue λ
- Thus, by convention eigenvectors are normalized to length 1 (multiplied by the reciprocal of the square root of the sum of squares of its components) and we have $|\mathbf{g}| = 1$
- Normalized eigenvectors are unique up to their sign

Example: Eigenvectors

- Plugging in $\lambda_1 = 1.8$ yields the adjunct matrix

$$adj(\mathbf{A} - \lambda_1 \mathbf{I}) = adj \begin{pmatrix} 1 - 1.8 & 0.8 \\ 0.8 & 1 - 1.8 \end{pmatrix} = \begin{pmatrix} -0.8 & -0.8 \\ -0.8 & -0.8 \end{pmatrix}$$

- Normalizing the first column provides

$$\frac{1}{\sqrt{0.8^2 + 0.8^2}} \begin{pmatrix} -0.8 \\ -0.8 \end{pmatrix} = \begin{pmatrix} -\sqrt{\frac{1}{2}} \\ -\sqrt{\frac{1}{2}} \end{pmatrix}$$

- Because eigenvectors are unique up to their sign, we consider as normalized eigenvector corresponding to λ_1

$$\mathbf{g}_1 = \begin{pmatrix} \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} \end{pmatrix}$$

Example: Eigenvectors

- Analogously, for eigenvalue $\lambda_2 = 0.2$ we get

$$adj(\mathbf{A} - \lambda_2 \mathbf{I}) = adj \begin{pmatrix} 1 - 0.2 & 0.8 \\ 0.8 & 1 - 0.2 \end{pmatrix} = \begin{pmatrix} 0.8 & -0.8 \\ -0.8 & 0.8 \end{pmatrix}$$

  (Notice that the two columns are linearly dependent)

- Normalizing the second column results in the eigenvector corresponding to λ_2

$$\mathbf{g}_2 = \begin{pmatrix} -\sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} \end{pmatrix}$$

Eigenvalues and Eigenvectors

Properties

- Eigenvalues $\lambda_1, \ldots, \lambda_n$ of an $(n \times n)$-matrix $\mathbf{A}$ have the following useful properties
  - $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i = \lambda_1 + \cdots + \lambda_n$
  - $\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i = \lambda_1 \cdot \cdots \cdot \lambda_n$
  - The rank of $\mathbf{A}$ equals the number of distinct non-zero eigenvalues of $\mathbf{A}$
  - A matrix $\mathbf{A}$ is called *positive semi-definite* (*positive definite*, respectively), which means that for all $\mathbf{x} \neq \mathbf{0}$ we have

    $$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0 \quad (\mathbf{x}'\mathbf{A}\mathbf{x} > 0, \textit{resp.})$$

    if for all eigenvalues λ of $\mathbf{A}$ we have $\lambda \geq 0$ ($\lambda > 0$, resp.)
  - The eigenvalues of a diagonal matrix are the diagonal elements

Eigenvalue Problem: Solution

- If all normalized eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$ of an $(n \times n)$-matrix $\mathbf{A}$ are collected in a matrix $\mathbf{G}$ then

$$\mathbf{G}'\mathbf{A}\mathbf{G} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_i & \ldots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & \lambda_n \end{pmatrix} = \mathbf{\Lambda}$$

(For the *orthogonal matrix* $\mathbf{G}$ we have $\mathbf{G}' = \mathbf{G}^{-1}$)

Eigenvalue Problem: Solution

- The matrix product $\mathbf{G}'\mathbf{A}\mathbf{G}$ thus results in a diagonal matrix, the diagonal elements of which are the eigenvalues of $\mathbf{A}$
- This provides

$$\mathrm{tr}(\mathbf{G}'\mathbf{A}\mathbf{G}) = \mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$$

and

$$\det(\mathbf{G}'\mathbf{A}\mathbf{G}) = \det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$$

Example

$$\mathbf{G}'\mathbf{A}\mathbf{G} = \begin{pmatrix} \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \\ -\sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} 1.8 & 0 \\ 0 & 0.2 \end{pmatrix}$$

Eigenvalues and -Vectors in R

Matrix

```
A <- matrix(c(1, 0.8, 0.8, 1), 2, 2)
```

Eigenvalues and eigenvectors

```
eig <- eigen(A)
lambda <- eig$values
G <- eig$vectors
```

Matrix diagonalization

```
Lambda <- t(G) %*% A %*% G

sum(diag(A))
sum(diag(Lambda))

det(A)
det(Lambda)
```

References

Books

Andres, J. (1996). Grundbegriffe der multivariaten Datenanalyse. In E. Erdfelder, R. Mausfeld, T. Meiser & G. Rudinger (Hrsg.), Handbuch Quantitative Methoden. Weinheim: Psychologie Verlags Union.

Fahrmeir, L., Hamerle, A. & Tutz, G. (1996). Multivariate statistische Verfahren. 2. Auflage. Berlin: De Gruyter. Anhang A

Meyberg, K. & Vachenauer, P. (2001). Höhere Mathematik 1. Berlin: Springer.

Stevens, J. (2001). Applied multivariate statistics for the social sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Online resources

`https://www.3blue1brown.com/`
`https://www.khanacademy.org/math/linear-algebra`
`https://open.umn.edu/opentextbooks/textbooks/5`
`https://textbooks.math.gatech.edu/ila/`