8. Linear mappings and matrices

A mapping f from IR^n to IR^m is called *linear* if it fulfills the following two properties:

(1) $f(\vec{a}+\vec{b}) = f(\vec{a}) + f(\vec{b})$ for all $\vec{a}, \vec{b} \in \mathrm{IR}^n$

(2) $f(\lambda\vec{a}) = \lambda f(\vec{a})$ for all $\lambda \in \mathrm{IR}$ and all $\vec{a} \in \mathrm{IR}^n$

Mappings of this sort appear frequently in the applications. E.g., some important geometrical mappings fall into the class of linear mappings: Rotations around the origin, reflections, projections, scalings, shear mappings...

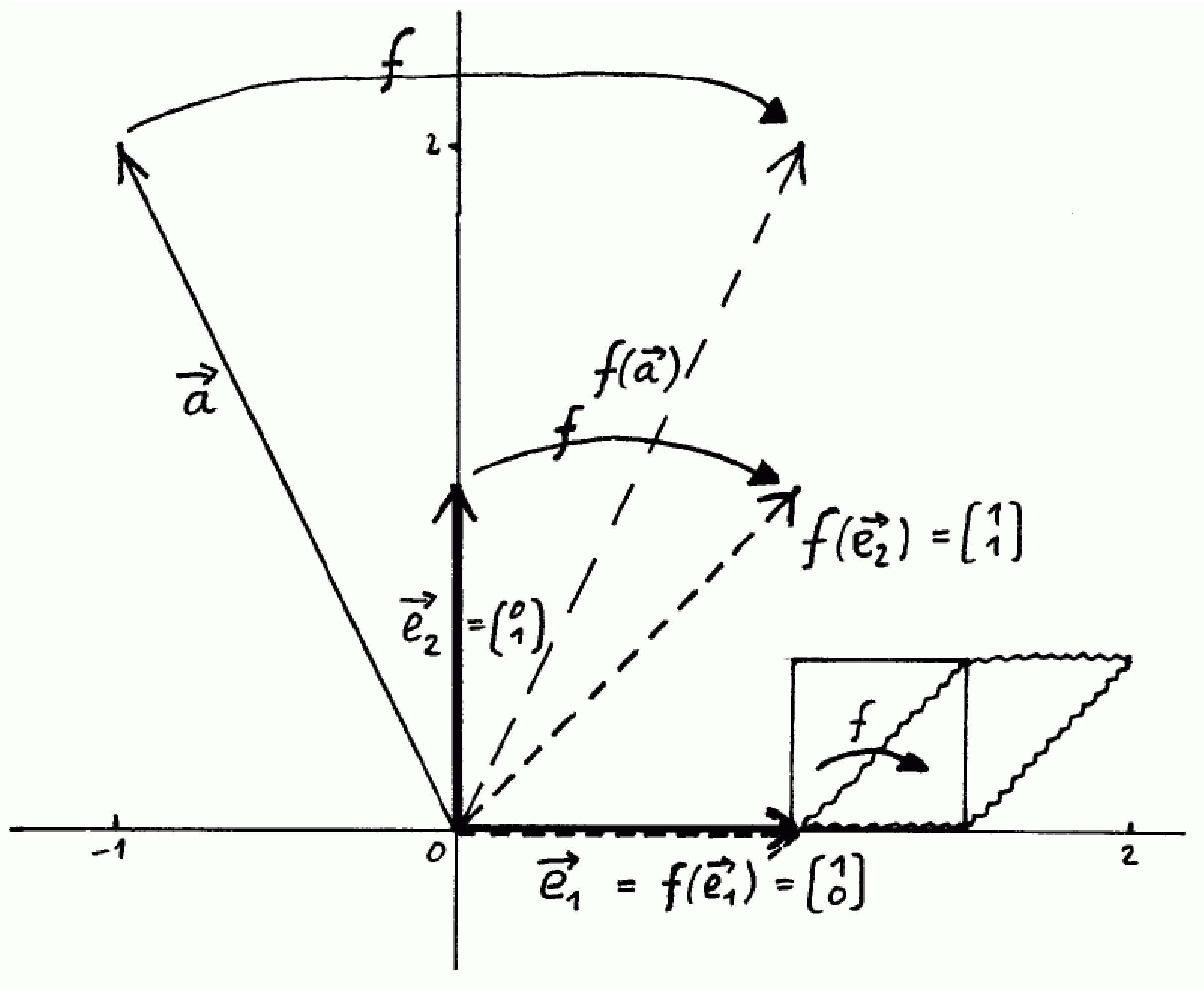
We show at the example of a shear mapping that such a mapping is completely determined (for all input vectors) if its effect on the vectors of the standard basis are known:

Example

Let f be the mapping from IR^2 to IR^2 which performs a *shear* along the *x* axis,

i.e., the image of each point under *f* can be found at the same height as the original point, but shifted along the *x* axis by a length which is proportional (in our example: even equal) to the *y* coordinate.

The figure illustrates the effect of f at the examples of the standard basis vectors and an arbitrary vector $\vec{a}$:



We have:

$$f : \mathbb{R}^2 \to \mathbb{R}^2$$

$$f\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}$$

$$f\begin{bmatrix}0\\2\end{bmatrix}=\begin{bmatrix}2\\2\end{bmatrix}=2\cdot f\begin{bmatrix}0\\1\end{bmatrix}$$

$$f\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}$$

f is indeed a linear mapping, that means:

$f(\vec{a}+\vec{b})=f(\vec{a})+f(\vec{b})$ and

$f(c\cdot\vec{a})=c\cdot f(\vec{a})$ are fulfilled.

The general formula for this shear mapping is apparently:

$$f\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix}$$

To get knowledge about the image $f\begin{bmatrix} x \\ y \end{bmatrix}$ of an arbitrary vector $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$, it is sufficient to know the *images of the vectors of the standard basis*, i.e., $f\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $f\begin{bmatrix} 0 \\ 1 \end{bmatrix}$:

$$\begin{bmatrix} x \\ y \end{bmatrix} = x \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x \cdot \vec{e}_1 + y \cdot \vec{e}_2$$

$$\Rightarrow \quad f\begin{bmatrix} x \\ y \end{bmatrix} = f(x \cdot \vec{e}_1 + y \cdot \vec{e}_2) \overset{f \text{ is linear}}{\downarrow\!=} x \cdot f(\vec{e}_1) + y \cdot f(\vec{e}_2)$$

Here: $f\begin{bmatrix} x \\ y \end{bmatrix} = x \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix}$,
confirming our formula above.

That means: These images, here $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, describe f completely.
They are put together in a *matrix*:

$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ = matrix of f.

In general:

Matrix of a linear mapping $f: \mathbb{R}^n \to \mathbb{R}^m$:

$$\begin{array}{cccc} \multicolumn{4}{c}{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}} \\ \uparrow & \uparrow & & \uparrow \\ \text{image of} & \text{image of} & & \text{image of} \\ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} & \cdots & \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \end{array}$$

has *m* rows and *n* columns

$\Rightarrow$ "matrix of type (*m*; *n*)"

all entries a_{ij} are real numbers

The matrix describes its associated linear mapping completely.

The result of the application of f to a vector $\vec{x} \in \mathbb{R}^n$ can easily be calculated as the *product* of the *matrix* of f *with the vector* $\vec{x}$.

In our example:

$$f\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \cdot x + 1 \cdot y \\ 0 \cdot x + 1 \cdot y \end{bmatrix} = \begin{bmatrix} x + y \\ y \end{bmatrix}$$

In the general case:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n \\ \vdots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \ldots + a_{mn} \cdot x_n \end{bmatrix}$$

Example: $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1\cdot 5 + 2\cdot 6 \\ 3\cdot 5 + 4\cdot 6 \end{bmatrix} = \begin{bmatrix} 17 \\ 39 \end{bmatrix}$

General definition of a matrix:

A *matrix* of type (m; n), also: $m \times n$ matrix ("m cross n"), is a system of $m \cdot n$ numbers a_{ij}, $i = 1, 2, ..., m$ and $j = 1, ..., n$, ordered in m rows and n columns:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

a_{ij} is called the *element* or *entry* of the i-th row and the j-th column. The $m \cdot n$ numbers a_{ij} are the *components* of the matrix.
A matrix of type (m; n) has m rows and n columns. Each row is an n-dimensional vector (row vector), and each column is an m-dimensional column vector.
The list of elements a_{ii} ($i = 1, 2, ..., r$ with $r = \min(m, n)$) is called the *principal diagonal* of the matrix.

Example:

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

A is of type (3; 4).

A has 3 row vectors:

$$\vec{z}_1 = (1, 4, -3, 2) \quad , \quad \vec{z}_2 = (2, 3, 0, -1) \quad , \quad \vec{z}_3 = (-3, 4, 1, 1)$$

and four column vectors:

$$\vec{s}_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} \quad , \quad \vec{s}_2 = \begin{bmatrix} 4 \\ 3 \\ 4 \end{bmatrix} \quad , \quad \vec{s}_3 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \quad , \quad \vec{s}_4 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

Its principal diagonal is 1; 3; 1.

Special forms of matrices:

• quadratic matrix:
If $m = n$, i.e., if the matrix *A* has as many rows as it has columns, *A* is called *quadratic*.

• $m = 1$: A matrix of type (1; n) is a row vector.

• $n = 1$: A matrix of type (m; 1) is a column vector.

• $m = n = 1$: A matrix of type (1; 1) can be identified with a single real number (i.e., its single entry).

• diagonal matrix:
If *A* is quadratic and all elements outside the principal diagonal are 0, *A* is called a *diagonal matrix*.

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

• unit matrix:
The unit matrix E is a diagonal matrix where all elements of the principal diagonal are 1.
It plays an important role: Its associated linear mapping is the *identical mapping* $f(\vec{x}) = \vec{x}$.

$$E = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

• zero matrix:
The matrix where *all* entries are 0 is called the zero matrix.

• triangular matrix:
A matrix where all elements below the principal diagonal are 0 is called an *upper triangular matrix*.

Example:

$$A = \begin{bmatrix} 5 & 2 & -1 & 7 \\ 0 & 3 & 1 & 5 \\ 0 & 0 & -1 & 10 \\ 0 & 0 & 0 & 42 \end{bmatrix}$$

Analogous: A matrix where all elements above the principal diagonal are 0 is called a *lower triangular matrix*.

Addition of matrices and *multiplication* of a matrix *with a scalar*:
These operations are defined in the same way as for vectors, i.e., component-wise.

Example:

$$5\cdot\begin{bmatrix}1 & 3\\0 & 2\end{bmatrix}+\begin{bmatrix}-1 & 0\\7 & 3\end{bmatrix}=\begin{bmatrix}5\cdot 1-1 & 5\cdot 3+0\\5\cdot 0+7 & 5\cdot 2+3\end{bmatrix}=\begin{bmatrix}4 & 15\\7 & 13\end{bmatrix}$$

Attention: Only matrices of the same type can be added.

Multiplication of a matrix with a column vector:
Defined as above, i.e.,

$$\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{m1} & \cdots & a_{mn}\end{bmatrix}\cdot\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = \begin{bmatrix}a_{11}\cdot x_1+a_{12}\cdot x_2+\ldots+a_{1n}\cdot x_n\\ \vdots\\ a_{m1}\cdot x_1+a_{m2}\cdot x_2+\ldots+a_{mn}\cdot x_n\end{bmatrix}.$$

The result corresponds to the image of the vector under the corresponding linear mapping.

Here, the matrix must have as many columns as the vector has components!

Transposition of a matrix:

Let A be a matrix of type (m; n). The matrix A^T of type (n; m), where its k-th row is the k-th column of A (k = 1, ..., m), is called the *transposed matrix* of A. (Transposition = reflection at the principal diagonal.)

$$\begin{bmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \ldots & a_{mn}\end{bmatrix}^T = \begin{bmatrix}a_{11} & a_{21} & \ldots & a_{m1}\\ a_{12} & a_{22} & \ldots & a_{m2}\\ \vdots & \vdots & & \vdots\\ a_{1n} & a_{2n} & \ldots & a_{mn}\end{bmatrix}$$

Example:

$$A = \begin{bmatrix} 1 & 4 \\ -2 & 3 \\ 3 & -1 \end{bmatrix}$$ of type (3; 2) $\Rightarrow$

$$A^T = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 3 & -1 \end{bmatrix}$$ of type (2; 3)

Special case: Transposition of a row vector (type (1; *m*)) gives a column vector (type (*m*; 1)), and vice versa.

Submatrix:

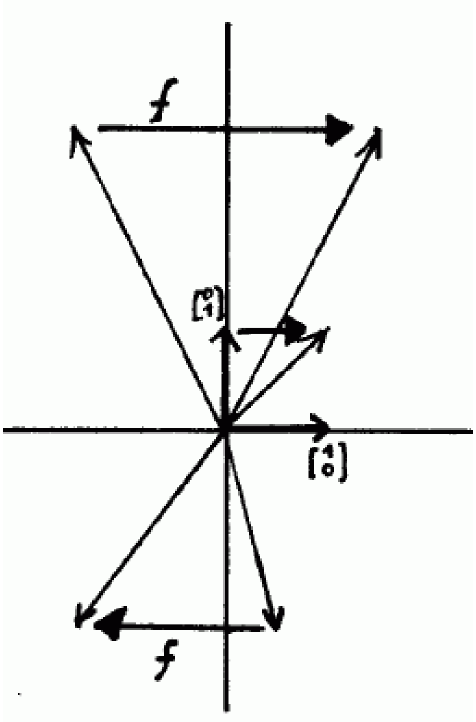
A *submatrix* of type (m–k; n–p) of a matrix A of type (m; n) is obtained by omitting k rows and p columns from A.
The special submatrix derived from A by omitting the i-th row and the j-th column is sometimes denoted A_{ij} .

We now come back to *linear mappings*, which were our entrance point to motivate the introduction of matrices.
Properties of linear mappings are reflected in numerical attributes of their corresponding matrices.

An important example is the so-called *rank* of a linear mapping.

We demonstrate it at two examples:

f : $IR^2 \rightarrow IR^2$ shear mapping (= example from above)



$$f\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}$$

$$f\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}$$

Matrix of f :

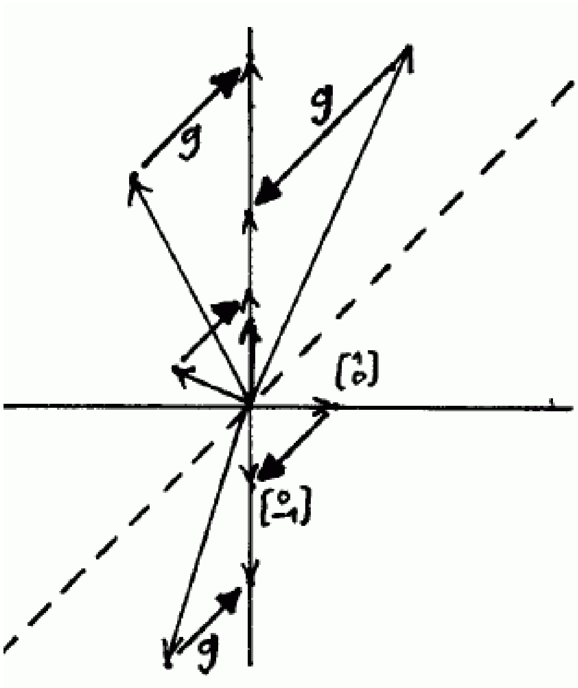
$$\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$$

The images

$$f\begin{bmatrix}1\\0\end{bmatrix}, f\begin{bmatrix}0\\1\end{bmatrix}$$

(i.e., the column vectors of the matrix of f)
are linearly independent,
they span the whole plane IR^2

g : $IR^2 \rightarrow IR^2$ projection along the principal diagonal onto the y axis



$$g\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}0\\-1\end{bmatrix}$$

$$g\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix}$$

Matrix of g :

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

The images

$$g\begin{bmatrix}1\\0\end{bmatrix}, g\begin{bmatrix}0\\1\end{bmatrix}$$

(i.e., the column vectors of the matrix of g)
are linearly dependent,
they are on the same line through 0 (i.e., on the y axis)

⇒ each vector is an image under f (f is surjective) rank f = 2 (= dimension of the plane)	⇒ only the y axis is the range of g (g is not surjective) rank g = 1 (= dimension of the line)

Definition:
The <u>rank of a matrix</u> A is the maximal number of linearly independent column vectors of A.
Notation: $rank(A)$, r(A).

This is consistent with our former definition:
$rank(A)$ = rank of the system of column vectors of A (as a vector system).

At the same time, it is the dimension of the range of the corresponding linear mapping of A.

Theorem:
$rank(A)$ is also the maximal number of linearly independent row vectors of A.

"column rank = row rank" !

Special cases:
The rank of the zero matrix is 0 (= smallest possible rank of a matrix).
The rank of E, the $n\times n$ unit matrix, is n (= largest possible rank of an $n\times n$ matrix).

The rank of an $m \times n$ matrix A can be at most the number of rows and at most the number of columns:
$0 \leq rank(A) \leq \min(m, n)$.

For determining the rank of a matrix, it is useful to know that under certain *elementary operations* the *rank* of a matrix *does not change*:

Elementary row operations

(1) Reordering of rows (particularly, switching of two rows)

(2) multiplication of a complete row by a number $c \neq 0$

(3) addition or omission of a row which is a linear combination of other rows

(4) addition of a linear combination of rows to another row.

Analogous for column operations.

Example:

$$A = \begin{bmatrix} 2 & 6 & -4 \\ 3 & 11 & 1 \\ -4 & -14 & 1 \end{bmatrix}$$

By applying elementary row operations, we transform *A* into an upper triangular matrix (parentheses are omitted for convenience):

$$\begin{array}{rrrl}
2 & 6 & -4 & :2 \\
3 & 11 & 1 & \\
-4 & -14 & 1 & \\
\hline
1 & 3 & -2 & \cdot 3 \quad \cdot 4 \\
3 & 11 & 1 & - \quad + \\
-4 & -14 & 1 & \\
\hline
1 & 3 & -2 & \\
0 & 2 & 7 & + \\
0 & -2 & -7 & \\
\hline
1 & 3 & -2 & \\
0 & 2 & 7 & \\
0 & 0 & 0 &
\end{array}$$

The rank of *A* must be the same as the rank of the matrix obtained in the end.

The rank of this triangular matrix can easily seen to be 2 (one zero row; zero rows are always linearly dependent! – The other two rows must be independent because of the first components 1 and 0.)