Matrix $\xrightarrow{\text{help us}}$ want to solve linear equations

3 Matrices and linear systems

table of numbers

Definition 3.1.

The set of all matrices with m rows and n columns *is notated as:*

$$\mathbb{R}^{m\times n} := \left\{ A = \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{\text{width } n} \Bigg\} \text{height } m \;:\; a_{ij} \in \mathbb{R}\,,\; i = 1, \dots, m,\; j = 1, \dots, n \right\}$$

$\mathbb{R}^{m\times n}$ gets an addition and scalar multiplication $\Big\}$ $\mathbb{R}^{m\times n}$ vector space

Definition 3.2. Matrix + Matrix = Matrix

Let $A, B \in \mathbb{R}^{m\times n}$*. The addition* $A+B \in \mathbb{R}^{m\times n}$ *is defined by*

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{A} + \underbrace{\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}}_{B} := \underbrace{\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}}_{A+B}.$$

new symbol

Example 3.3.

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{\in \mathbb{R}^{2\times 2}} + \underbrace{\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}}_{\in \mathbb{R}^{2\times 2}} = \begin{pmatrix} 1+1 & 2+0 \\ 3+2 & 4-1 \end{pmatrix} = \underline{\begin{pmatrix} 2 & 2 \\ 5 & 3 \end{pmatrix}} \in \mathbb{R}^{2\times 2}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 7 & 8 \\ 9 & 10 \end{pmatrix} \text{ is } \underline{\text{not}} \text{ defined}$$

Attention!

The addition $A + B$ is only defined for matrices with the same height and the same width.

Definition 3.4. Scalar · Matrix = Matrix

Let $A \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$. Then the scalar multiplication $\lambda \cdot A \in \mathbb{R}^{m\times n}$ is defined by:

$$\lambda\cdot \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{A} := \underbrace{\begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}}_{\lambda\cdot A}.$$

new symbol

Example 3.5.

$$2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2\cdot 1 & 2\cdot 2 \\ 2\cdot 3 & 2\cdot 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}. = A + A$$

$\mathbb{R}^{m\times n}$ with $+$ and $\lambda\cdot$ is a vector space like $\mathbb{R}^n$
(same calculation rules, e.g., distributive law)

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$ zero matrix = neutral element w.r.t. $+$.

3.1 Introduction to systems of linear equations

We start with some easy examples:

Xavier is two years older than Yasmin.	$x - y = 2$
Together they are 40 years old.	$x + y = 40$
How old is Xavier and how old is Yasmin?	$x = ?, \quad y = ?$

How to solve it? ~ 2 equations 2 unknowns

This was an example with two unknowns (x and y). Here we give an example for three unknowns. (x, y and z):

$$\begin{array}{rrrcr} 2x & -3y & +4z & = & -7 \\ -3x & +y & -z & = & 0 \\ 20x & +10y & & = & 80 \\ & 10y & +25z & = & 90 \end{array}$$

4 equations
3 unknowns

$\rightsquigarrow$ $x = ?$, $y = ?$, $z = ?$

Unknowns: $(x_1, \dots, x_n)$, $n \in \mathbb{N}$

Number of equations: m, $m \in \mathbb{N}$

Question: Which values for $x_1, \dots, x_n$ satisfy all equations at once.

↳ a solution

Here, the most important part is that the equations are linear. The exact definition will follow later. The sloppy way to say that an equation is linear is:

power is 1

$$\text{constant} \cdot x_1 + \text{constant} \cdot x_2 + \cdots + \text{constant} \cdot x_n = \text{constant}\,. \tag{3.1}$$

just addition

As you can see, there are a lot of constants that have to be numeric.

Definition 3.6. System of linear equations (LES)

Let $m, n \in \mathbb{N}$ be two natural numbers. A system of linear equations or a linear equation system (abbreviation: LES) with m equations and n unknowns $x_1, x_2, \dots, x_n$ is given by:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\} \tag{LES}$$

Here, a_{ij} and b_i are given numbers, mostly just real numbers. A solution of the LES is a choice of values for $x_1, \dots, x_n$ such that all m equations are satisfied.

$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$

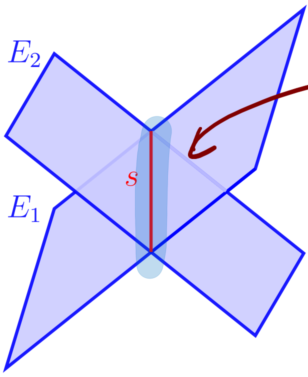
vector

1st question: Existence?

2nd question: Uniqueness?

Example 3.7. Having three unknowns x_1, x_2, x_3, we could have different cases for the set of solutions:

$\mathbb{R}^3 \rightsquigarrow$ Equations describe affine subspaces (e.g. translated plane)

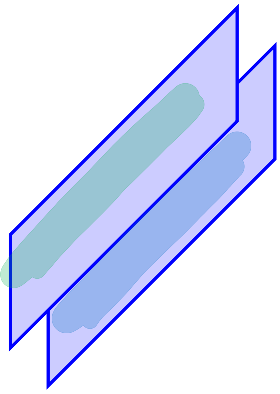


solution set

existence ✓

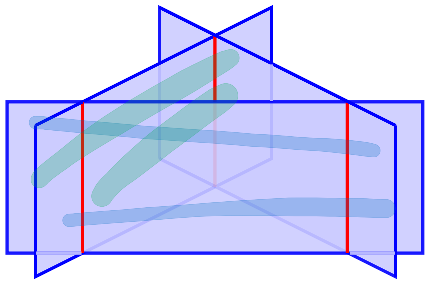
uniqueness ↯

no solution



2 equations ↯

or



no solution

3 equations ↯

Definition 3.8. LES in matrix notation

Let $A \in \mathbb{R}^{m\times n}$ *with entries* $a_{ij} \in \mathbb{R}$ *and* $\mathbf{b} \in \mathbb{R}^m$ *with entries* $b_i \in \mathbb{R}$*. Then*

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

represents (LES) *from above, where* $\mathbf{x} \in \mathbb{R}^n$.

The two examples from above in this notation:

$$\underbrace{\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}}_{\text{Matrix}} \cdot \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\text{Vector}} = \underbrace{\begin{pmatrix} 2 \\ 40 \end{pmatrix}}_{\text{Vector}}, \quad \begin{pmatrix} 2 & -3 & 4 \\ -3 & 1 & -1 \\ 20 & 10 & 0 \\ 0 & 10 & 25 \end{pmatrix} \underline{\begin{pmatrix} x \\ y \\ z \end{pmatrix}} = \underline{\begin{pmatrix} -7 \\ 0 \\ 80 \\ 90 \end{pmatrix}}.$$

Matrix - Vector multiplication!

Definition 3.9. Matrix · Vector = Vector

Let $m, n \in \mathbb{N}$ and

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

The product $A\mathbf{x} = A \cdot \mathbf{x}$ (where we mostly do not use a dot) is given as the vector

$$A\mathbf{x} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \in \mathbb{R}^m.$$

Attention!

The width of A has to be the same as the height of $\mathbf{x}$! Otherwise $A\mathbf{x}$ is not defined.

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 0 + 3 \cdot 1 \\ 4 \cdot 1 + 5 \cdot 0 + 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \end{pmatrix}$$

3.2 Some words about matrices

For a matrix $A \in \mathbb{R}^{m \times n}$ the number m is called the number of rows and n the number of columns. The matrix A is a rectangle with height m and width n.

As special cases, we note:

- $A \in \mathbb{R}^{n \times n}$ (i.e. $m = n$) is called a square matrix or quadratic matrix
- $A \in \mathbb{R}^{m \times 1}$ is a column vector of size m $\begin{pmatrix} 4 \\ 10 \end{pmatrix}$
- $A \in \mathbb{R}^{1 \times n}$ is a row vector of size n $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$
- $A \in \mathbb{R}^{1 \times 1}$ is a scalar, just a real number.

$A = (2) \iff A$ is a scalar

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

diagonal of matrix a_{ii}

$$A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

diagonal matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

upper triangle matrix

3.3 Looking at the columns and the associated linear map

One way to imagine a matrix in $\mathbb{R}^{m\times n}$ is to see it as a collection of n columns of size m:

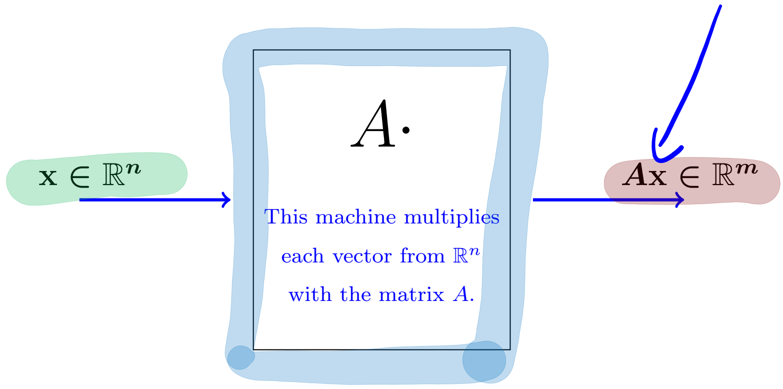
$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} | & & | \\ \mathbf{a}_1 & \dots & \mathbf{a}_n \\ | & & | \end{pmatrix}, \text{ where } \mathbf{a}_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \in \mathbb{R}^m$$

$A\mathbf{x}$ is a linear combination of the columns of A

$$A\mathbf{x} = \begin{pmatrix} | & & | \\ \mathbf{a}_1 & \dots & \mathbf{a}_n \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} | \\ \mathbf{a}_1 \\ | \end{pmatrix} + \dots + x_n \begin{pmatrix} | \\ \mathbf{a}_n \\ | \end{pmatrix} \tag{3.2}$$

scalar factors

linear combination of the columns of A.



$f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$

The function f_A defined by the matrix A

$$f_A : \mathbb{R}^n \to \mathbb{R}^m, \qquad \textit{with} \qquad f_A : \mathbf{x} \mapsto A\mathbf{x} \tag{3.3}$$

$\leadsto$ see later: f_A linear map

3.4 Looking at the rows

Above, we have considered a matrix $A \in \mathbb{R}^{m\times n}$ as a collection of columns and defined a linear map $f_A : \mathbb{R}^n \to \mathbb{R}^m$. However, we may also see A as a collection of m rows of size n:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} -\ \boldsymbol{\alpha}_1^T\ - \\ \cdots \\ -\ \boldsymbol{\alpha}_m^T\ - \end{pmatrix}, \quad \text{where } \boldsymbol{\alpha}_i^T = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$$

Here, we use the notation T for the transpose of a column vector. The result is a row vector with the same entries. We fix this as a space:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^T = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$$

$$\mathbb{R}^{1\times n} = \{\mathbf{x}^T = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{R}\}$$

Since a row vector $\mathbf{u}^T \in \mathbb{R}^{1\times n}$ is just a very flat matrix, the product with a column vector $\mathbf{v} \in \mathbb{R}^n$ is well-defined:

$$\mathbf{u}^T\mathbf{v} = (u_1v_1 + \cdots + u_nv_n) \in \mathbb{R}^{1\times 1}.$$

$u, v \in \mathbb{R}^n$

$u^T \in \mathbb{R}^{1\times n}$

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 2 + 8 + 18 \end{pmatrix} = \begin{pmatrix} 28 \end{pmatrix} \in \mathbb{R}^1$$

$= 28$

$$\langle u, v\rangle = u^T v$$

Standard inner product in each row

$A\mathbf{x}$ is the scalar product of $\mathbf{x}$ with the rows of A

$$A\mathbf{x} = \begin{pmatrix} -\ \boldsymbol{\alpha}_1^T\ - \\ \ddots \\ -\ \boldsymbol{\alpha}_m^T\ - \end{pmatrix} \begin{pmatrix} | \\ \mathbf{x} \\ | \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_1^T\mathbf{x} \\ \vdots \\ \boldsymbol{\alpha}_n^T\mathbf{x} \end{pmatrix} \tag{3.4}$$

3.5 Matrix multiplication

$A \cdot B$

Let $A \in \mathbb{R}^{m\times n}$.

$b \in \mathbb{R}^n \rightsquigarrow Ab \in \mathbb{R}^m$

$b_1, \dots, b_k \in \mathbb{R}^n \rightsquigarrow Ab_1, \dots, Ab_k \in \mathbb{R}^m$

$$B = \begin{pmatrix} | & & | \\ b_1 & \cdots & b_k \\ | & & | \end{pmatrix} \in \mathbb{R}^{n\times k} \rightsquigarrow \begin{pmatrix} | & & | \\ Ab_1 & \cdots & Ab_k \\ | & & | \end{pmatrix} \in \mathbb{R}^{m\times k}$$

$$A \begin{pmatrix} | & & | \\ \mathbf{b}_1 & \dots & \mathbf{b}_k \\ | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k \\ | & | & & | \end{pmatrix} \quad \text{for } k \text{ column vectors } \mathbf{b}_1, \dots, \mathbf{b}_k\,.$$

The result is a matrix with m rows and k columns and denoted by AB. It is called the *matrix product of A and B*.

Definition 3.10. Matrix product

For matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$, the matrix product is defined as

$$AB = \begin{pmatrix} - \boldsymbol{\alpha}_1^T - \\ \ddots \\ - \boldsymbol{\alpha}_m^T - \end{pmatrix} \begin{pmatrix} | & & | \\ \mathbf{b}_1 & \dots & \mathbf{b}_k \\ | & & | \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_1^T \mathbf{b}_1 & \cdots & \boldsymbol{\alpha}_1^T \mathbf{b}_k \\ \vdots & & \vdots \\ \boldsymbol{\alpha}_m^T \mathbf{b}_1 & \cdots & \boldsymbol{\alpha}_1^T \mathbf{b}_k \end{pmatrix} \in \mathbb{R}^{m\times k}. \tag{3.5}$$

Or in other words: AB is the $m \times k$-matrix that has the following entries:

$$(AB)_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$$

for $i = 1, \dots, m$ and $j = 1, \dots, k$.

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}}_{A \in \mathbb{R}^{2\times 3}} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}}_{B \in \mathbb{R}^{3\times 2}} = \underbrace{\begin{pmatrix} 4 & 5 \\ 10 & 11 \end{pmatrix}}_{\mathbb{R}^{2\times 2}}$$

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

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 10 & 11 \end{pmatrix}$$

$$\cdot : \mathbb{R}^{m\times n} \times \mathbb{R}^{n\times k} \longrightarrow \mathbb{R}^{m\times k}$$

Attention!

The product AB is only defined if the width of A coincides with the height of B. The "inner dimensions" have to match.

Special cases:

- $A = \mathbf{a}^T \in \mathbb{R}^{1\times n}$, $B = \mathbf{b} \in \mathbb{R}^{n\times 1}$: $AB = \mathbf{a}^T\mathbf{b} \in \mathbb{R}^{1\times 1}$
- $A = \mathbf{a} \in \mathbb{R}^{n\times 1}$, $B = \mathbf{b}^T \in \mathbb{R}^{1\times m}$: $AB = \mathbf{a}\mathbf{b}^T \in \mathbb{R}^{n\times m}$, $(AB)_{ij} = a_i b_j$ *(called rank 1 matrix)*

Example 3.11. Just calculate some examples:

(a) We combine the following matrix dimensions $(2 \times 2) \cdot (2 \times 3) \Rightarrow 2 \times 3$:

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}}_{B} = \left(\underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 \\ 4 \end{pmatrix}}_{\mathbf{b}_1} \quad \underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 2 \\ 5 \end{pmatrix}}_{\mathbf{b}_2} \quad \underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 3 \\ 6 \end{pmatrix}}_{\mathbf{b}_3} \right)$$

2x3

$$= \begin{pmatrix} \underbrace{\begin{matrix}1\cdot 1+2\cdot 4\\ 3\cdot 1+4\cdot 4\end{matrix}}_{A\mathbf{b}_1} & \underbrace{\begin{matrix}1\cdot 2+2\cdot 5\\ 3\cdot 2+4\cdot 5\end{matrix}}_{A\mathbf{b}_2} & \underbrace{\begin{matrix}1\cdot 3+2\cdot 6\\ 3\cdot 3+4\cdot 6\end{matrix}}_{A\mathbf{b}_3} \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{pmatrix}$$

$$\begin{matrix} & \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} & \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{pmatrix} \end{matrix}$$

(b) Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. But: $B \cdot A$ is defined ✓

$$\begin{matrix} & \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} & \end{matrix}$$

AB is not defined

$$\begin{matrix} & \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} & \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \end{matrix}$$

(c) Now the matrix dimensions $(3 \times 1) \cdot (1 \times 3) \Rightarrow 3 \times 3$:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} (1\ 2\ 3) = \left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \underbrace{1}_{\mathbf{b}_1} \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \underbrace{2}_{\mathbf{b}_2} \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \underbrace{3}_{\mathbf{b}_3} \right) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$$

↪ rank one matrix

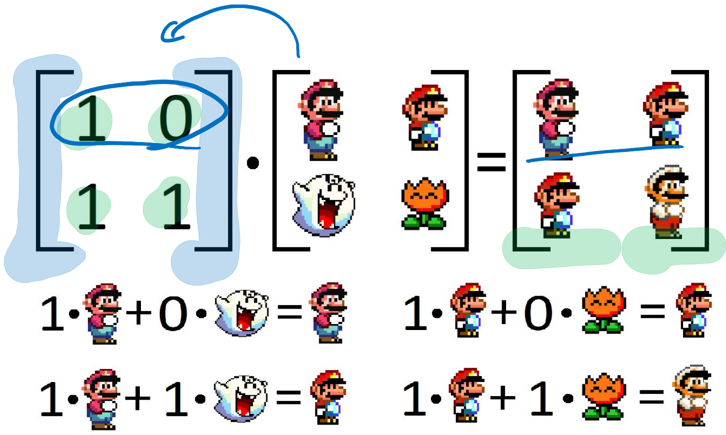
$$\begin{matrix} & (1\ 2\ 3) \\ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix} \end{matrix}$$

(d) Now the matrix dimensions $(1 \times 3) \cdot (3 \times 1) \Rightarrow 1 \times 1$:

$$(1\ 2\ 3) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = (1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3) = (14) = 14$$

identification

(e) A 2×2-example:



nothing changes in the first row

(Source of the picture: g33ktheory.com)

We can also ask what happens if we multiply a row vector $\mathbf{x}^T$ from the left to a matrix B. By definition, we get:

linear combinations of the rows.

$$\mathbf{x}^T B = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} - \beta_1^T - \\ \vdots \\ - \beta_n^T - \end{pmatrix} = x_1 \left(- \beta_1^T - \right) + \cdots + x_n \left(- \beta_n^T - \right) .$$

This means the product $\mathbf{x}^T B$ is a linear combination of the rows of B. This is an analogy that $A\mathbf{x}$ is a linear combination of the columns of A, cf. equation (3.4).

Remark:

Now we can see the matrix product as introduced

$$AB = \begin{pmatrix} | & | & & | \\ A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k \\ | & | & & | \end{pmatrix}$$

This means that each column of AB consists of a linear combination of the columns from A.
Seeing the product with the other eye

$$AB = \begin{pmatrix} - \boldsymbol{\alpha}_1^T - \\ \vdots \\ - \boldsymbol{\alpha}_m^T - \end{pmatrix} B = \begin{pmatrix} - \boldsymbol{\alpha}_1^T B - \\ \vdots \\ - \boldsymbol{\alpha}_m^T B - \end{pmatrix} ,$$

we see that each row of AB consists of a linear combination of the rows from B.

Now, we summarise the properties of the matrix multiplication.

Proposition 3.12. Properties of the matrix product

(a) *For all* $A, B \in \mathbb{R}^{m\times n}$ *and* $C \in \mathbb{R}^{n\times k}$ *and* $D \in \mathbb{R}^{\ell\times m}$ *we have:*

$$(A+B)\cdot C = A\cdot C + B\cdot C \quad \text{and} \quad D\cdot(A+B) = D\cdot A + D\cdot B.$$

(b) *For all* $A \in \mathbb{R}^{m\times n}$ *and* $B \in \mathbb{R}^{n\times k}$ *and* $\lambda \in \mathbb{R}$ *we have:*

$$\lambda\cdot(A\cdot B) = (\lambda\cdot A)\cdot B = A\cdot(\lambda\cdot B).$$

(c) *Associative rule: For all* $A \in \mathbb{R}^{m\times n}$ *and* $B \in \mathbb{R}^{n\times k}$ *and* $C \in \mathbb{R}^{k\times\ell}$ *we have:*

$$A\cdot(B\cdot C) = (A\cdot B)\cdot C.$$

Proof. All these rules follow from the definition of the matrix product of A and B,

$$(AB)_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}\,,$$

real numbers

and the fact that these rules hold for the real numbers $a_{ir}, b_{rj} \in \mathbb{R}$. For example, for showing (c):

$$(A(BC))_{ij} = \sum_{r=1}^{n} a_{ir}(BC)_{rj} = \sum_{r=1}^{n} a_{ir}\sum_{z=1}^{n} b_{rz}c_{zj} = \sum_{z=1}^{n}\left(\sum_{r=1}^{n} a_{ir}b_{rz}\right)c_{zj} = ((AB)C)_{ij}\,.$$

using real number rules

Properties (a) and (b) are left as an exercise. □

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

$$\neq$$

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

Attention! No commutative rule

In general, we have for two matrices:

$$AB \neq BA \quad \text{(in general)}.$$

Remark:

We thus have the following interpretations of the matrix vector product AB*:*

- *the columns of* B *are used to build linear combinations of the columns of* A*,*
- *the rows of* A *are used to build linear combinations of the rows of* B*,*
- *each row* $\boldsymbol{\alpha}_i^T$ *of* A *and each column* $\mathbf{b}_j$ *of* B *are multiplied to form an entry of the product:* $(AB)_{ij} = \boldsymbol{\alpha}_i^T\mathbf{b}_j$*,*
- *each column* $\mathbf{a}_i$ *of* A *and each row* $\boldsymbol{\beta}_i^T$ *of* B *is combined to a rank-1 matrix* $\mathbf{a}_i\boldsymbol{\beta}_i^T$*, and the matrices are added up,*

All these interpretations are equally valid, and from situation to situation, we can change our point of view to gain additional insights.