Linear Algebra

Deisenroth, Chapter 2

a.fahim@ut.ac.ir

ناسازگار

$$+\begin{cases} x_1 + x_2 + x_3 = 3 & (1) \\ x_1 - x_2 + 2x_3 = 2 & (2) \end{cases}$$
$$2x_1 + 3x_3 = 1 \quad (3)$$

$\rightarrow 2x_1 + 3x_3 = 5$

$(1, 1, 1)$

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 && (1) \\ -\; x_1 - x_2 + 2x_3 &= 2 && (2) \\ x_2 + x_3 &= 2 && (3) \end{aligned}$$

$x_2 = 1$

$x_2 = 2 - x_3$

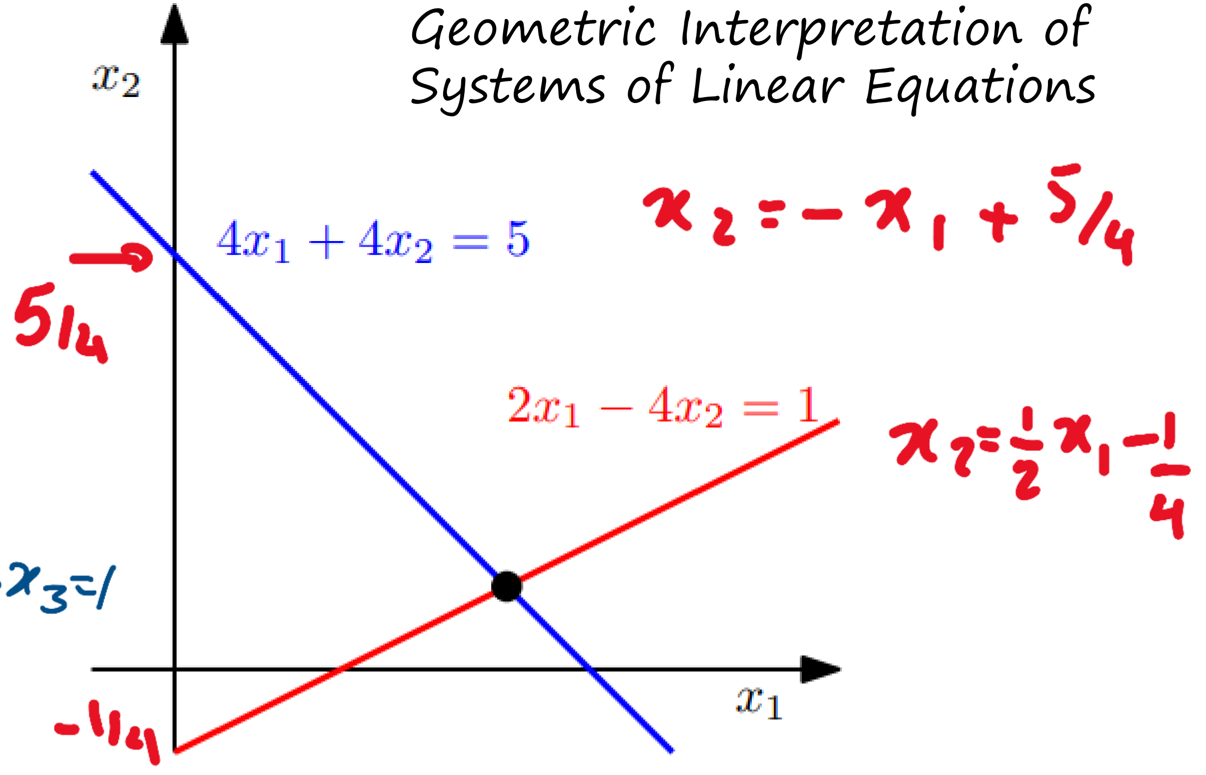
$0 + 2x_2 + 3x_3 = 5 \rightarrow 4 - 2x_3 + 3x_3 = 5 \Rightarrow x_3 = 1$

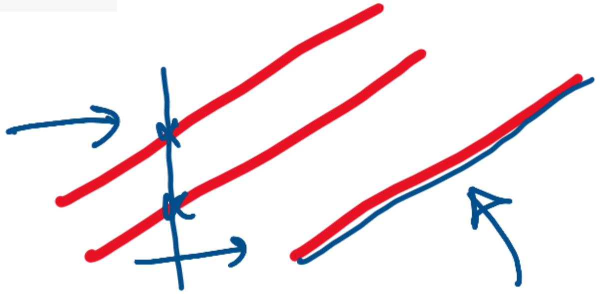
$$+\begin{cases} x_1 + x_2 + x_3 = 3 & (1) \\ x_1 - x_2 + 2x_3 = 2 & (2) \end{cases}$$
$$2x_1 + 3x_3 = 5 \quad (3)$$

$2x_1 + 3x_3 = 5$

بی‌نهایت

Geometric Interpretation of Systems of Linear Equations





$$\begin{cases} 4x_1 + 4x_2 = 5 \\ 2x_1 - 4x_2 = 1 \end{cases}$$

a.fahim@ut.ac.ir

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

m سطر (معادله)

n مجهول

$$\iff \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 1 \\ 4x_1 - 2x_2 - 7x_3 &= 8 \\ 9x_1 + 5x_2 - 3x_3 &= 2 \end{aligned}$$

$$\underbrace{\begin{bmatrix} 2 & 3 & 5 \\ 4 & -2 & -7 \\ 9 & 5 & -3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} = \underbrace{\begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix}}_{B}.$$

Definition 2.1 (Matrix). With $m, n \in \mathbb{N}$ a real-valued (m, n) *matrix* $\boldsymbol{A}$ is an $m \cdot n$-tuple of elements a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, which is ordered according to a rectangular scheme consisting of m rows and n columns:

بردار سطری

$$[a_{11} \quad a_{12} \quad \cdots \quad a_{1n}]$$

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \underbrace{a_{ij} \in \mathbb{R}}_{a_{ij}} . \tag{2.11}$$

بردار ستونی

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

By convention $(1, n)$-matrices are called *rows* and $(m, 1)$-matrices are called *columns*. These special matrices are also called *row/column vectors*.

Matrix Addition and Multiplication

The sum of two matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{m\times n}$ is defined as the element-wise sum, i.e.,

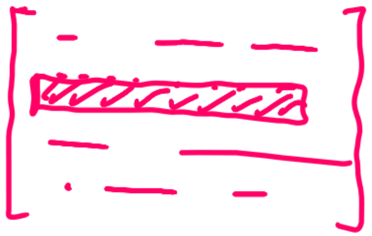
$$\boldsymbol{A}+\boldsymbol{B} := \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}\,. \tag{2.12}$$

For matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times k}$, the elements c_{ij} of the product $\boldsymbol{C} = \boldsymbol{AB} \in \mathbb{R}^{m\times k}$ are computed as

$$c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}, \qquad i = 1,\ldots,m, \quad j = 1,\ldots,k. \tag{2.13}$$

$n=3$

$c_{25} = a_{21}b_{15} + a_{22}b_{25} + a_{23}b_{35}$

$A =$ 

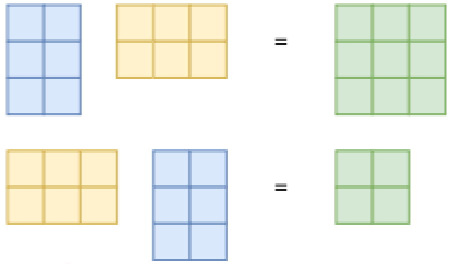
$B =$ 

For $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 3}$, $\boldsymbol{B} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2}$, we obtain

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 2 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

$$\boldsymbol{BA} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 & 2 \\ -2 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Figure 2.5 Even if both matrix multiplications $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are defined, the dimensions of the results can be different.



- *Associativity:*

$$\forall \boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{B} \in \mathbb{R}^{n\times p}, \boldsymbol{C} \in \mathbb{R}^{p\times q} : (\boldsymbol{AB})\boldsymbol{C} = \underbrace{\boldsymbol{A}(\boldsymbol{BC})}_{m\times q} \qquad (2.18)$$

- *Distributivity:*

$$\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m\times n}, \boldsymbol{C}, \boldsymbol{D} \in \mathbb{R}^{n\times p} : (\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC} \qquad (2.19a)$$

$$\boldsymbol{A}(\boldsymbol{C}+\boldsymbol{D}) = \boldsymbol{AC} + \boldsymbol{AD} \qquad (2.19b)$$

- Multiplication with the identity matrix:

$$\forall \boldsymbol{A} \in \mathbb{R}^{m\times n} : \boldsymbol{I}_m \boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A} \qquad (2.20)$$

Note that $\boldsymbol{I}_m \neq \boldsymbol{I}_n$ for $m \neq n$.

$$\boldsymbol{I}_n := \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

Definition 2.3 (Inverse). Consider a square matrix $A \in \mathbb{R}^{n\times n}$. Let matrix $B \in \mathbb{R}^{n\times n}$ have the property that $AB = I_n = BA$.

B is called the *inverse* of A and denoted by A^{-1}.

Unfortunately, not every matrix A possesses an inverse A^{-1}.

If this inverse does exist, A is called *regular/invertible/nonsingular*, otherwise *singular*.

$$A := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2} \qquad \rightarrow \qquad A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

if and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Furthermore, we can generally use the determinant to check whether a matrix is invertible.

a.fahim@ut.ac.ir

Definition 2.4 (Transpose). For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ the matrix $\boldsymbol{B} \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$ is called the *transpose* of $\boldsymbol{A}$. We write $\boldsymbol{B} = \boldsymbol{A}^\top$.

In general, $\boldsymbol{A}^\top$ can be obtained by writing the columns of $\boldsymbol{A}$ as the rows of $\boldsymbol{A}^\top$. The following are important properties of inverses and transposes:

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I} = \boldsymbol{A}^{-1}\boldsymbol{A} \tag{2.26}$$

$$(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \tag{2.27}$$

$$(\boldsymbol{A}+\boldsymbol{B})^{-1} \neq \boldsymbol{A}^{-1} + \boldsymbol{B}^{-1} \tag{2.28}$$

$$(\boldsymbol{A}^\top)^\top = \boldsymbol{A} \tag{2.29}$$

$$(\boldsymbol{A}\boldsymbol{B})^\top = \boldsymbol{B}^\top\boldsymbol{A}^\top \tag{2.30}$$

$$(\boldsymbol{A}+\boldsymbol{B})^\top = \boldsymbol{A}^\top + \boldsymbol{B}^\top \tag{2.31}$$

Definition 2.5 (Symmetric Matrix). A matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is *symmetric* if $\boldsymbol{A} = \boldsymbol{A}^\top$.

Remark (Sum and Product of Symmetric Matrices). The sum of symmetric matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times n}$ is always symmetric. However, although their product is always defined, it is generally not symmetric:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}. \tag{2.32}$$

Multiplication by a Scalar

Let us look at what happens to matrices when they are multiplied by a scalar $\lambda \in \mathbb{R}$. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$. Then $\lambda\boldsymbol{A} = \boldsymbol{K}$, $K_{ij} = \lambda\, a_{ij}$. Practically, λ scales each element of $\boldsymbol{A}$. For $\lambda, \psi \in \mathbb{R}$, the following holds:

- *Associativity:*
  $(\lambda\psi)\boldsymbol{C} = \lambda(\psi\boldsymbol{C}), \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$
- $\lambda(\boldsymbol{B}\boldsymbol{C}) = (\lambda\boldsymbol{B})\boldsymbol{C} = \boldsymbol{B}(\lambda\boldsymbol{C}) = (\boldsymbol{B}\boldsymbol{C})\lambda, \quad \boldsymbol{B} \in \mathbb{R}^{m\times n}, \boldsymbol{C} \in \mathbb{R}^{n\times k}$.
  Note that this allows us to move scalar values around.
- $(\lambda\boldsymbol{C})^\top = \boldsymbol{C}^\top\lambda^\top = \boldsymbol{C}^\top\lambda = \lambda\boldsymbol{C}^\top$ since $\lambda = \lambda^\top$ for all $\lambda \in \mathbb{R}$.
- *Distributivity:*
  $(\lambda + \psi)\boldsymbol{C} = \lambda\boldsymbol{C} + \psi\boldsymbol{C}, \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$
  $\lambda(\boldsymbol{B} + \boldsymbol{C}) = \lambda\boldsymbol{B} + \lambda\boldsymbol{C}, \quad \boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{m\times n}$

Solving Systems of Linear Equations

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

$$a_{11}x_1 + \cdots + a_{1n}x_n = b_1$$

$$\vdots$$

$$a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \,,$$

$$\begin{bmatrix} 1 & 0 & 8 & -4 \\ 0 & 1 & 2 & 12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 42 \\ 8 \end{bmatrix}$$

$$\underbrace{\sum_{i=1}^{4} x_i \boldsymbol{c}_i = \boldsymbol{b},}$$

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 8 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 12 \end{bmatrix} =$$

$$42 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 8 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Particular and General Solution

$$\boldsymbol{b} = \begin{bmatrix} 42 \\ 8 \end{bmatrix} = 42 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 8 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Therefore, a solution is $[42, 8, 0, 0]^\top$.

This solution is called a particular solution special solution.

Particular and General Solution

1. Find a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$.
2. Find all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$.
3. Combine the solutions from steps 1. and 2. to the general solution.

Neither the general nor the particular solution is unique. ◇

2 معادله

۴ مجهول

$$\begin{bmatrix} 1 & 0 & 8 & -4 \\ 0 & 1 & 2 & 12 \end{bmatrix} \left(\lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix} \right) = \lambda_1 (8\boldsymbol{c}_1 + 2\boldsymbol{c}_2 - \boldsymbol{c}_3) = \boldsymbol{0}\,.$$

$$\left\{ \boldsymbol{x} \in \mathbb{R}^4 : \boldsymbol{x} = \begin{bmatrix} 42 \\ 8 \\ 0 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} -4 \\ 12 \\ 0 \\ -1 \end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}.$$

Elementary Transformations

- Exchange of two equations (rows in the matrix representing the system of equations)
- Multiplication of an equation (row) with a constant $\lambda \in \mathbb{R}\backslash\{0\}$
- Addition of two equations (rows)

$$\begin{bmatrix} 0 & 1 & \overset{0}{\cancel{x}} & 5 & 1 \\ 0 & 0 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

A matrix is in **row echelon form** when it satisfies the following conditions.
1 – The first non-zero element in each row, called the leading entry, is 1.
2 – Each leading entry is in a column to the right of the leading entry in the previous row.
3 – Rows with all zero elements, if any, are below rows having a non-zero element.

A matrix is in **reduced row echelon form** when it satisfies the following conditions.
1 – The matrix is in row echelon form (i.e., it satisfies the three conditions listed above).
2 – The leading entry in each row is the only non-zero entry in its column.

For $a \in \mathbb{R}$, we seek all solutions of the following system of equations:

$$\begin{array}{rcrcrcrcrcr} -2x_1 & + & 4x_2 & - & 2x_3 & - & x_4 & + & 4x_5 & = & -3 \\ 4x_1 & - & 8x_2 & + & 3x_3 & - & 3x_4 & + & x_5 & = & 2 \\ x_1 & - & 2x_2 & + & x_3 & - & x_4 & + & x_5 & = & 0 \\ x_1 & - & 2x_2 & & & - & 3x_4 & + & 4x_5 & = & a \end{array}. \tag{2.44}$$

We start by converting this system of equations into the compact matrix notation $\boldsymbol{Ax} = \boldsymbol{b}$. We no longer mention the variables $\boldsymbol{x}$ explicitly and build the *augmented matrix* (in the form $[\boldsymbol{A} \mid \boldsymbol{b}]$)

$$\left[\begin{array}{rrrrr|r} -2 & 4 & -2 & -1 & 4 & -3 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ 1 & -2 & 1 & -1 & 1 & 0 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right] \begin{array}{l} \text{Swap with } R_3 \\ \\ \text{Swap with } R_1 \\ \\ \end{array}$$

Remark (Basic and Free Variables). The variables corresponding to the pivots in the row-echelon form are called *basic variables* and the other variables are *free variables*. For example, in (2.45), x_1, x_3, x_4 are basic variables, whereas x_2, x_5 are free variables. ◇

In the previous example, we would try to find $\lambda_1, \lambda_2, \lambda_3$ so that

$$\lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}. \tag{2.48}$$

From here, we find relatively directly that $\lambda_3 = 1, \lambda_2 = -1, \lambda_1 = 2$. When we put everything together, we must not forget the non-pivot columns for which we set the coefficients implicitly to 0. Therefore, we get the particular solution $\boldsymbol{x} = [2, 0, -1, 1, 0]^\top$. ◇

This (augmented) matrix is in a convenient form, the *row-echelon form* (REF). Reverting this compact notation back into the explicit notation with the variables we seek, we obtain

$$\begin{array}{rcrcrcrcrcr} x_1 & - & \cancel{2x_2} & + & \cancel{x_3} & - & x_4 & + & \cancel{x_5} & = & 0 \\ & & & & x_3 & - & x_4 & + & \cancel{3x_5} & = & -2 \\ & & & & & & x_4 & - & 2x_5 & = & 1 \\ & & & & & & & & 0 & = & a+1 \end{array}. \tag{2.45}$$

$x_1 = 2$ -1

Only for $a = -1$ this system can be solved. A *particular solution* is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}. \tag{2.46}$$

The *general solution*, which captures the set of all possible solutions, is

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}. \tag{2.47}$$

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

$A\vec{x}_1 = \vec{0}$

$A(\vec{x} + \lambda_1 \vec{x}_1 + \lambda_2 \vec{x}_2) = \vec{b} + \lambda_1 \vec{0} + \lambda_2 \vec{0}$

a.fahim@ut.ac.ir

Example 2.7 (Reduced Row Echelon Form)
Verify that the following matrix is in reduced row-echelon form (the pivots are in **bold**):

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{1} & 3 & 0 & 0 & 3 \\ 0 & 0 & \mathbf{1} & 0 & 9 \\ 0 & 0 & 0 & \mathbf{1} & -4 \end{bmatrix}. \tag{2.49}$$

The key idea for finding the solutions of $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ is to look at the *non-pivot columns*, which we will need to express as a (linear) combination of the pivot columns. The reduced row echelon form makes this relatively straightforward, and we express the non-pivot columns in terms of sums and multiples of the pivot columns that are on their left: The second column is 3 times the first column (we can ignore the pivot columns on the right of the second column). Therefore, to obtain $\mathbf{0}$, we need to subtract the second column from three times the first column. Now, we look at the fifth column, which is our second non-pivot column. The fifth column can be expressed as 3 times the first pivot column, 9 times the second pivot column, and -4 times the third pivot column. We need to keep track of the indices of the pivot columns and translate this into 3 times the first column, 0 times the second column (which is a non-pivot column), 9 times the third column (which is our second pivot column), and -4 times the fourth column (which is the third pivot column). Then we need to subtract the fifth column to obtain $\mathbf{0}$. In the end, we are still solving a homogeneous equation system.

To summarize, all solutions of $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}, \boldsymbol{x} \in \mathbb{R}^5$ are given by

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}. \tag{2.50}$$

The Minus-1 Trick

Example 2.8 (Minus-1 Trick)
Let us revisit the matrix in (2.49), which is already in reduced REF:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \tag{2.53}$$

We now augment this matrix to a 5×5 matrix by adding rows of the form (2.52) at the places where the pivots on the diagonal are missing and obtain

$Ax = 0$

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & \mathbf{-1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & \mathbf{-1} \end{bmatrix} . \tag{2.54}$$

From this form, we can immediately read out the solutions of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ by taking the columns of $\tilde{\boldsymbol{A}}$, which contain -1 on the diagonal:

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix} , \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\} , \tag{2.55}$$

which is identical to the solution in (2.50) that we obtained by "insight".

Calculating the Inverse

$[A]_{n\times n}$ $Ax=b$ $\underbrace{x=A^{-1}b}$

To compute the inverse $\boldsymbol{A}^{-1}$ of $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, we need to find a matrix $\boldsymbol{X}$ that satisfies $\boldsymbol{AX} = \boldsymbol{I}_n$. Then, $\boldsymbol{X} = \boldsymbol{A}^{-1}$. We can write this down as a set of simultaneous linear equations $\boldsymbol{AX} = \boldsymbol{I}_n$, where we solve for $\boldsymbol{X} = [\boldsymbol{x}_1|\cdots|\boldsymbol{x}_n]$. We use the augmented matrix notation for a compact representation of this set of systems of linear equations and obtain

$$\underbrace{[\boldsymbol{A}|\boldsymbol{I}_n]}_{n\times 2n} \quad \rightsquigarrow \cdots \rightsquigarrow \quad [\boldsymbol{I}_n|\boldsymbol{A}^{-1}]\,. \tag{2.56}$$

Example 2.9 (Calculating an Inverse Matrix by Gaussian Elimination)
To determine the inverse of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{2.57}$$

we write down the augmented matrix

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

and use Gaussian elimination to bring it into reduced row-echelon form

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 2 & -2 & 2 \\ 0 & 1 & 0 & 0 & 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 2 \end{array}\right],$$

such that the desired inverse is given as its right-hand side:

$$\boldsymbol{A}^{-1} = \begin{bmatrix} -1 & 2 & -2 & 2 \\ 1 & -1 & 2 & -2 \\ 1 & -1 & 1 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}. \tag{2.58}$$

We can verify that (2.58) is indeed the inverse by performing the multiplication $\boldsymbol{A}\boldsymbol{A}^{-1}$ and observing that we recover $\boldsymbol{I}_4$.

Algorithms for Solving a System of Linear Equations

$AX = b \quad X = A^{-1}b$

$$\underset{m\times n}{A}x = b \iff \underset{n\times m}{A^\top}\,\underset{m\times n}{A}x = A^\top b \iff x = \underbrace{(A^\top A)^{-1}A^\top}b$$

$\underset{n\times n}{[A^\top A]}$

Moore-Penrose pseudo-inverse

In practice, systems of many linear equations are solved indirectly, by either stationary iterative methods, such as the Richardson method, the Jacobi method, the ~~Gauß-Seidel~~ method, and the successive over-relaxation method, or Krylov subspace methods, such as conjugate gradients, generalized minimal residual, or biconjugate gradients. We refer to the books by Stoer and Burlirsch (2002), Strang (2003), and Liesen and Mehrmann (2015) for further details.

Let x_* be a solution of $Ax = b$. The key idea of these iterative methods is to set up an iteration of the form

$$x^{(k+1)} = Cx^{(k)} + d \tag{2.60}$$

a.fahim@ut.ac.ir

Groups

Definition 2.7 (Group). Consider a set $\mathcal{G}$ and an operation $\otimes : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ defined on $\mathcal{G}$. Then $G := (\mathcal{G}, \otimes)$ is called a *group* if the following hold:

1. *Closure* of $\mathcal{G}$ under $\otimes$: $\forall x, y \in \mathcal{G} : x \otimes y \in \mathcal{G}$
2. *Associativity:* $\forall x, y, z \in \mathcal{G} : (x \otimes y) \otimes z = x \otimes (y \otimes z)$
3. *Neutral element:* $\exists e \in \mathcal{G}\ \forall x \in \mathcal{G} : x \otimes e = x$ and $e \otimes x = x$
4. *Inverse element:* $\forall x \in \mathcal{G}\ \exists y \in \mathcal{G} : x \otimes y = e$ and $y \otimes x = e$, where e is the neutral element. We often write x^{-1} to denote the inverse element of x.

If additionally $\forall x, y \in \mathcal{G} : x \otimes y = y \otimes x$, then $G = (\mathcal{G}, \otimes)$ is an *Abelian group* (commutative).

Definition 2.8 (General Linear Group). The set of regular (invertible) matrices $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is a group with respect to matrix multiplication as defined in (2.13) and is called *general linear group* $GL(n, \mathbb{R})$. However, since matrix multiplication is not commutative, the group is not Abelian.

Example 2.10 (Groups)
Let us have a look at some examples of sets with associated operations and see whether they are groups:

- $(\mathbb{Z}, +)$ is an Abelian group.
- $(\mathbb{N}_0, +)$ is not a group: Although $(\mathbb{N}_0, +)$ possesses a neutral element (0), the inverse elements are missing.
- $(\mathbb{Z}, \cdot)$ is not a group: Although $(\mathbb{Z}, \cdot)$ contains a neutral element (1), the inverse elements for any $z \in \mathbb{Z}, z \neq \pm 1$, are missing.
- $(\mathbb{R}, \cdot)$ is not a group since 0 does not possess an inverse element.
- $(\mathbb{R}\backslash\{0\}, \cdot)$ is Abelian.
- $(\mathbb{R}^n, +), (\mathbb{Z}^n, +), n \in \mathbb{N}$ are Abelian if $+$ is defined componentwise, i.e.,

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n). \tag{2.61}$$

  Then, $(x_1, \cdots, x_n)^{-1} := (-x_1, \cdots, -x_n)$ is the inverse element and $e = (0, \cdots, 0)$ is the neutral element.
- $(\mathbb{R}^{m \times n}, +)$, the set of $m \times n$-matrices is Abelian (with componentwise addition as defined in (2.61)).

- Let us have a closer look at $(\mathbb{R}^{n \times n}, \cdot)$, i.e., the set of $n \times n$-matrices with matrix multiplication as defined in (2.13).
  - Closure and associativity follow directly from the definition of matrix multiplication.
  - Neutral element: The identity matrix $\boldsymbol{I}_n$ is the neutral element with respect to matrix multiplication "$\cdot$" in $(\mathbb{R}^{n \times n}, \cdot)$.
  - Inverse element: If the inverse exists ($\boldsymbol{A}$ is regular), then $\boldsymbol{A}^{-1}$ is the inverse element of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, and in exactly this case $(\mathbb{R}^{n \times n}, \cdot)$ is a group, called the *general linear group*.

Definition 2.9 (Vector Space). A real-valued *vector space* $V = (\mathcal{V}, +, \cdot)$ is a set $\mathcal{V}$ with two operations

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V} \tag{2.62}$$

$$\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V} \tag{2.63}$$

where

1. $(\mathcal{V}, +)$ is an Abelian group
2. Distributivity:
   1. $\forall \lambda \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V} : \lambda \cdot (\boldsymbol{x} + \boldsymbol{y}) = \lambda \cdot \boldsymbol{x} + \lambda \cdot \boldsymbol{y}$
   2. $\forall \lambda, \psi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V} : (\lambda + \psi) \cdot \boldsymbol{x} = \lambda \cdot \boldsymbol{x} + \psi \cdot \boldsymbol{x}$
3. Associativity (outer operation): $\forall \lambda, \psi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V} : \lambda \cdot (\psi \cdot \boldsymbol{x}) = (\lambda\psi) \cdot \boldsymbol{x}$
4. Neutral element with respect to the outer operation: $\forall \boldsymbol{x} \in \mathcal{V} : 1 \cdot \boldsymbol{x} = \boldsymbol{x}$

The elements $\boldsymbol{x} \in V$ are called *vectors*. The neutral element of $(\mathcal{V}, +)$ is the zero vector $\mathbf{0} = [0, \ldots, 0]^\top$, and the inner operation $+$ is called *vector addition*. The elements $\lambda \in \mathbb{R}$ are called *scalars* and the outer operation $\cdot$ is a *multiplication by scalars*. Note that a scalar product is something different, and we will get to this in Section 3.2.

Example 2.11 (Vector Spaces)
Let us have a look at some important examples:

- $\mathcal{V} = \mathbb{R}^n, n \in \mathbb{N}$ is a vector space with operations defined as follows:
  - Addition: $\boldsymbol{x}+\boldsymbol{y} = (x_1, \ldots, x_n)+(y_1, \ldots, y_n) = (x_1+y_1, \ldots, x_n+y_n)$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$
  - Multiplication by scalars: $\lambda\boldsymbol{x} = \lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$ for all $\lambda \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^n$
- $\mathcal{V} = \mathbb{R}^{m\times n}, m, n \in \mathbb{N}$ is a vector space with
  - Addition: $\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$ is defined elementwise for all $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{V}$
  - Multiplication by scalars: $\lambda\boldsymbol{A} = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}$ as defined in Section 2.2. Remember that $\mathbb{R}^{m\times n}$ is equivalent to $\mathbb{R}^{mn}$.
- $\mathcal{V} = \mathbb{C}$, with the standard definition of addition of complex numbers.

Remark. The vector spaces $\mathbb{R}^n, \mathbb{R}^{n\times 1}, \mathbb{R}^{1\times n}$ are only different in the way we write vectors. In the following, we will not make a distinction between $\mathbb{R}^n$ and $\mathbb{R}^{n\times 1}$, which allows us to write n-tuples as *column vectors*

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}. \tag{2.64}$$

Definition 2.10 (Vector Subspace). Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{U} \subseteq \mathcal{V}$, $\mathcal{U} \neq \emptyset$. Then $U = (\mathcal{U}, +, \cdot)$ is called *vector subspace* of V (or *linear subspace*) if U is a vector space with the vector space operations $+$ and $\cdot$ restricted to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$. We write $U \subseteq V$ to denote a subspace U of V.

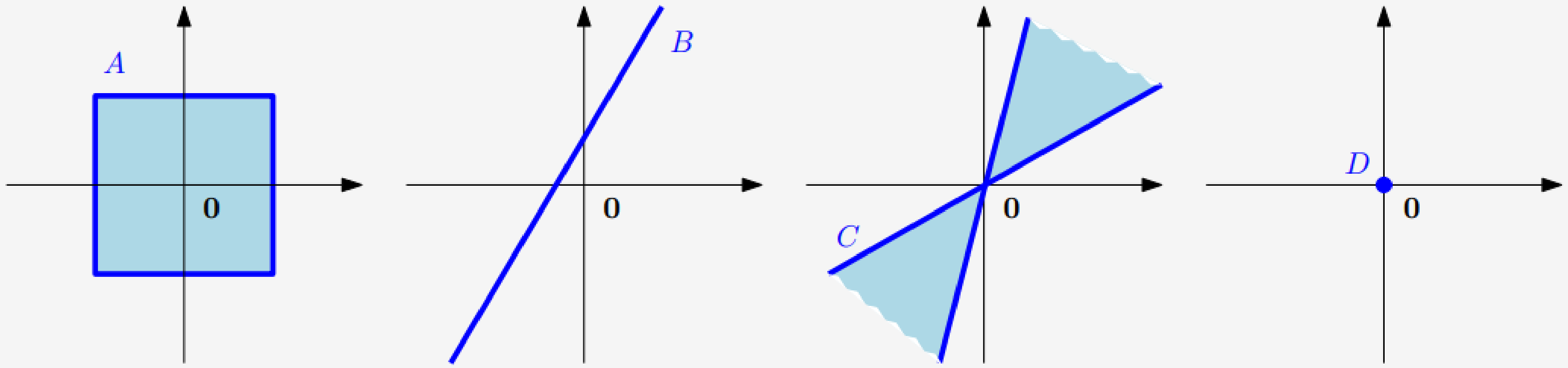
If $\mathcal{U} \subseteq \mathcal{V}$ and V is a vector space, then U naturally inherits many properties directly from V because they hold for all $\boldsymbol{x} \in \mathcal{V}$, and in particular for all $\boldsymbol{x} \in \mathcal{U} \subseteq \mathcal{V}$. This includes the Abelian group properties, the distributivity, the associativity and the neutral element. To determine whether $(\mathcal{U}, +, \cdot)$ is a subspace of V we still do need to show

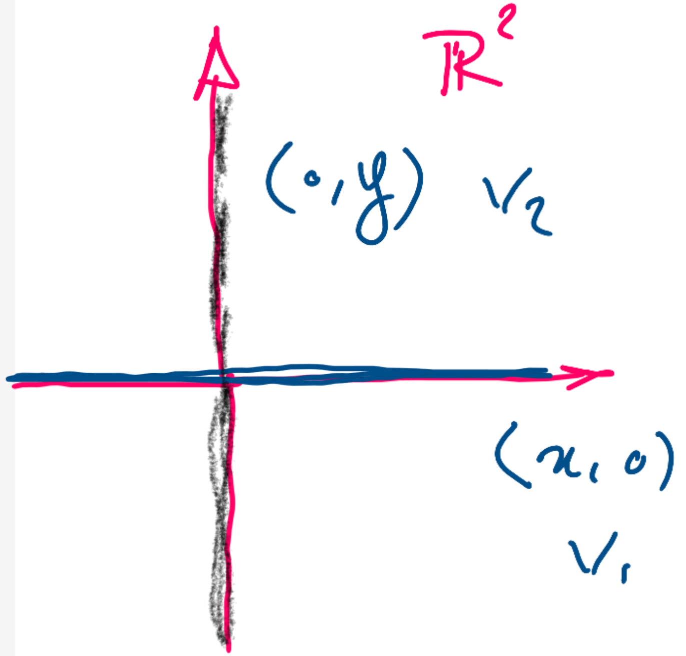
1. $\mathcal{U} \neq \emptyset$, in particular: $\mathbf{0} \in \mathcal{U}$
2. Closure of U:
   a. With respect to the outer operation: $\forall \lambda \in \mathbb{R} \, \forall \boldsymbol{x} \in \mathcal{U} : \lambda \boldsymbol{x} \in \mathcal{U}$.
   b. With respect to the inner operation: $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{U} : \boldsymbol{x} + \boldsymbol{y} \in \mathcal{U}$.

a.fahim@ut.ac.ir

Example 2.12 (Vector Subspaces)
Let us have a look at some examples:

- For every vector space V, the trivial subspaces are V itself and $\{\mathbf{0}\}$.
- Only example D in Figure 2.6 is a subspace of $\mathbb{R}^2$ (with the usual inner/outer operations). In A and C, the closure property is violated; B does not contain $\mathbf{0}$.
- The solution set of a homogeneous system of linear equations $\boldsymbol{Ax} = \mathbf{0}$ with n unknowns $\boldsymbol{x} = [x_1, \ldots, x_n]^\top$ is a subspace of $\mathbb{R}^n$.
- The solution of an inhomogeneous system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$, $\boldsymbol{b} \neq \mathbf{0}$ is not a subspace of $\mathbb{R}^n$.
- The intersection of arbitrarily many subspaces is a subspace itself.

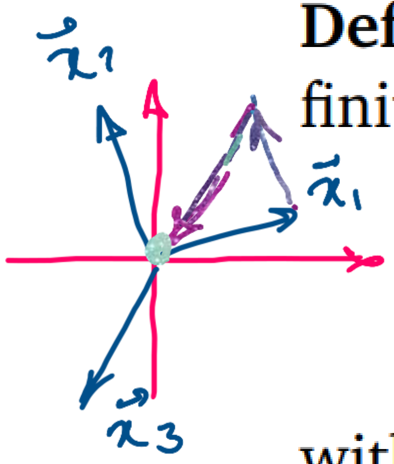




$V_1 \cap V_2 = \vec{0}$

Remark. Every subspace $U \subseteq (\mathbb{R}^n, +, \cdot)$ is the solution space of a homogeneous system of linear equations $\boldsymbol{Ax} = \mathbf{0}$ for $\boldsymbol{x} \in \mathbb{R}^n$. ◇

a.fahim@ut.ac.ir



Definition 2.11 (Linear Combination). Consider a vector space V and a finite number of vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in V$. Then, every $\boldsymbol{v} \in V$ of the form

$$\boldsymbol{v} = \lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i \in V \tag{2.65}$$

with $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ is a *linear combination* of the vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$.

$$\vec{0} = \sum_{i=1}^{k} \lambda_i \vec{x}_i \qquad \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$$

$$\lambda_i \neq 0$$

Definition 2.12 (Linear (In)dependence). Let us consider a vector space V with $k \in \mathbb{N}$ and $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in V$. If there is a non-trivial linear combination, such that $\boldsymbol{0} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$ with at least one $\lambda_i \neq 0$, the vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$ are *linearly dependent*. If only the trivial solution exists, i.e., $\lambda_1 = \dots = \lambda_k = 0$ the vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$ are *linearly independent*.

- A practical way of checking whether vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$ are linearly independent is to use Gaussian elimination: Write all vectors as columns of a matrix $\boldsymbol{A}$ and perform Gaussian elimination until the matrix is in row echelon form (the reduced row-echelon form is unnecessary here):

$A = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_k \\ | & | & & | \end{bmatrix}$

  - The pivot columns indicate the vectors, which are linearly independent of the vectors on the left. Note that there is an ordering of vectors when the matrix is built.
  - The non-pivot columns can be expressed as linear combinations of the pivot columns on their left. For instance, the row-echelon form

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{2.66}$$

    tells us that the first and third columns are pivot columns. The second column is a non-pivot column because it is three times the first column.

Example 2.14
Consider $\mathbb{R}^4$ with

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \quad \boldsymbol{x}_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix}. \tag{2.67}$$

To check whether they are linearly dependent, we follow the general approach and solve

$$\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \lambda_3 \boldsymbol{x}_3 = \lambda_1 \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + \lambda_3 \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix} = \boldsymbol{0} \tag{2.68}$$

$\lambda_1 = \lambda_2 = \lambda_3 = 0$

for $\lambda_1, \ldots, \lambda_3$. We write the vectors $\boldsymbol{x}_i$, $i = 1, 2, 3$, as the columns of a matrix and apply elementary row operations until we identify the pivot columns:

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.69}$$

Here, every column of the matrix is a pivot column. Therefore, there is no non-trivial solution, and we require $\lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0$ to solve the equation system. Hence, the vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ are linearly independent.

Example 2.15
Consider a set of linearly independent vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4 \in \mathbb{R}^n$ and

$$\begin{array}{rclclclcl} \boldsymbol{x}_1 & = & \boldsymbol{b}_1 & - & 2\boldsymbol{b}_2 & + & \boldsymbol{b}_3 & - & \boldsymbol{b}_4 \\ \boldsymbol{x}_2 & = & -4\boldsymbol{b}_1 & - & 2\boldsymbol{b}_2 & & & + & 4\boldsymbol{b}_4 \\ \boldsymbol{x}_3 & = & 2\boldsymbol{b}_1 & + & 3\boldsymbol{b}_2 & - & \boldsymbol{b}_3 & - & 3\boldsymbol{b}_4 \\ \boldsymbol{x}_4 & = & 17\boldsymbol{b}_1 & - & 10\boldsymbol{b}_2 & + & 11\boldsymbol{b}_3 & + & \boldsymbol{b}_4 \end{array} . \tag{2.73}$$

Are the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_4 \in \mathbb{R}^n$ linearly independent? To answer this question, we investigate whether the column vectors

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -4 \\ -2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 17 \\ -10 \\ 11 \\ 1 \end{bmatrix} \right\} \tag{2.74}$$

are linearly independent. The reduced row-echelon form of the corresponding linear equation system with coefficient matrix

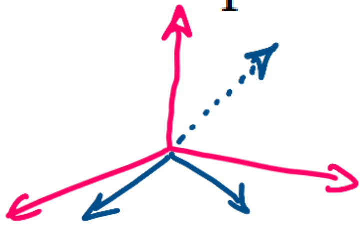
$$\boldsymbol{A} = \begin{bmatrix} 1 & -4 & 2 & 17 \\ -2 & -2 & 3 & -10 \\ 1 & 0 & -1 & 11 \\ -1 & 4 & -3 & 1 \end{bmatrix} \tag{2.75}$$

is given as

$$\begin{bmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -15 \\ 0 & 0 & 1 & -18 \\ 0 & 0 & 0 & 0 \end{bmatrix} . \tag{2.76}$$

We see that the corresponding linear equation system is non-trivially solvable: The last column is not a pivot column, and $\boldsymbol{x}_4 = -7\boldsymbol{x}_1 - 15\boldsymbol{x}_2 - 18\boldsymbol{x}_3$. Therefore, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_4$ are linearly dependent as $\boldsymbol{x}_4$ can be expressed as a linear combination of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_3$.

Definition 2.13 (Generating Set and Span). Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and set of vectors $\mathcal{A} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subseteq \mathcal{V}$. If every vector $\boldsymbol{v} \in \mathcal{V}$ can be expressed as a linear combination of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$, $\mathcal{A}$ is called a *generating set* of V. The set of all linear combinations of vectors in $\mathcal{A}$ is called the *span* of $\mathcal{A}$. If $\mathcal{A}$ spans the vector space V, we write $V = \text{span}[\mathcal{A}]$ or $V = \text{span}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k]$.



Definition 2.14 (Basis). Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and $\mathcal{A} \subseteq \mathcal{V}$. A generating set $\mathcal{A}$ of V is called *minimal* if there exists no smaller set $\tilde{\mathcal{A}} \subsetneq \mathcal{A} \subseteq \mathcal{V}$ that spans V. Every linearly independent generating set of V is minimal and is called a *basis* of V.

Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{B} \subseteq \mathcal{V}, \mathcal{B} \neq \emptyset$. Then, the following statements are equivalent:

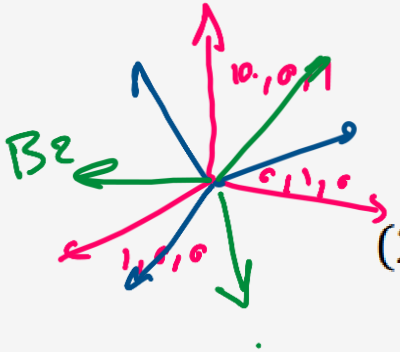
- $\mathcal{B}$ is a basis of V.
- $\mathcal{B}$ is a minimal generating set.
- $\mathcal{B}$ is a maximal linearly independent set of vectors in V, i.e., adding any other vector to this set will make it linearly dependent.
- Every vector $\boldsymbol{x} \in V$ is a linear combination of vectors from $\mathcal{B}$, and every linear combination is unique, i.e., with

$$\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{b}_i = \sum_{i=1}^{k} \psi_i \boldsymbol{b}_i \tag{2.77}$$

and $\lambda_i, \psi_i \in \mathbb{R}$, $\boldsymbol{b}_i \in \mathcal{B}$ it follows that $\lambda_i = \psi_i,\ i = 1, \ldots, k$.

- In $\mathbb{R}^3$, the *canonical/standard basis* is

$$\mathcal{B} = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\}. \tag{2.78}$$



- Different bases in $\mathbb{R}^3$ are

$$\mathcal{B}_1 = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix} \right\}, \mathcal{B}_2 = \left\{ \begin{bmatrix}0.5\\0.8\\0.4\end{bmatrix}, \begin{bmatrix}1.8\\0.3\\0.3\end{bmatrix}, \begin{bmatrix}-2.2\\-1.3\\3.5\end{bmatrix} \right\}. \tag{2.79}$$

- The set

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

$$\mathcal{A} = \left\{ \begin{bmatrix}1\\2\\3\\4\end{bmatrix}, \begin{bmatrix}2\\-1\\0\\2\end{bmatrix}, \begin{bmatrix}1\\1\\0\\-4\end{bmatrix} \right\} \tag{2.80}$$

is linearly independent, but not a generating set (and no basis) of $\mathbb{R}^4$: For instance, the vector $[1, 0, 0, 0]^\top$ cannot be obtained by a linear combination of elements in $\mathcal{A}$.

Remark. Every vector space V possesses a basis $\mathcal{B}$. The preceding examples show that there can be many bases of a vector space V, i.e., there is no unique basis. However, all bases possess the same number of elements, the *basis vectors*. ◊

We only consider finite-dimensional vector spaces V. In this case, the *dimension* of V is the number of basis vectors of V, and we write $\dim(V)$. If $U \subseteq V$ is a subspace of V, then $\dim(U) \leqslant \dim(V)$ and $\dim(U) = \dim(V)$ if and only if $U = V$. Intuitively, the dimension of a vector space can be thought of as the number of independent directions in this vector space.

Remark. A basis of a subspace $U = \text{span}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_m] \subseteq \mathbb{R}^n$ can be found by executing the following steps:

1. Write the spanning vectors as columns of a matrix $\boldsymbol{A}$
2. Determine the row-echelon form of $\boldsymbol{A}$.
3. The spanning vectors associated with the pivot columns are a basis of U.

$$A = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_m \\ | & | & & | \end{bmatrix}$$

Example 2.17 (Determining a Basis)
For a vector subspace $U \subseteq \mathbb{R}^5$, spanned by the vectors

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 2 \\ -2 \end{bmatrix}, \quad \boldsymbol{x}_3 = \begin{bmatrix} 3 \\ -4 \\ 3 \\ 5 \\ -3 \end{bmatrix}, \quad \boldsymbol{x}_4 = \begin{bmatrix} -1 \\ 8 \\ -5 \\ -6 \\ 1 \end{bmatrix} \in \mathbb{R}^5, \tag{2.81}$$

we are interested in finding out which vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_4$ are a basis for U. For this, we need to check whether $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_4$ are linearly independent. Therefore, we need to solve

$$\sum_{i=1}^{4} \lambda_i \boldsymbol{x}_i = \boldsymbol{0}, \tag{2.82}$$

which leads to a homogeneous system of equations with matrix

A

$$[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4] = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{bmatrix}. \tag{2.83}$$

With the basic transformation rules for systems of linear equations, we obtain the row-echelon form

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the pivot columns indicate which set of vectors is linearly independent, we see from the row-echelon form that $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_4$ are linearly independent (because the system of linear equations $\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \lambda_4 \boldsymbol{x}_4 = \boldsymbol{0}$ can only be solved with $\lambda_1 = \lambda_2 = \lambda_4 = 0$). Therefore, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_4\}$ is a basis of U.

$x_1 \quad x_2 \quad x_4$

استقلال خطی دارند

$\lambda_1 x_1 + \lambda_2 x_2 + \lambda_4 x_4 = 0$

$\lambda_1 = \lambda_2 = \lambda_4 = 0$

The number of linearly independent columns of a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ equals the number of linearly independent rows and is called the *rank* of $\boldsymbol{A}$ and is denoted by $\mathrm{rk}(\boldsymbol{A})$.

Remark. The rank of a matrix has some important properties:

- $\mathrm{rk}(\boldsymbol{A}) = \mathrm{rk}(\boldsymbol{A}^\top)$, i.e., the column rank equals the row rank.
- The columns of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ span a subspace $U \subseteq \mathbb{R}^m$ with $\dim(U) = \mathrm{rk}(\boldsymbol{A})$. Later we will call this subspace the *image* or *range*. A basis of U can be found by applying Gaussian elimination to $\boldsymbol{A}$ to identify the pivot columns.
- The rows of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ span a subspace $W \subseteq \mathbb{R}^n$ with $\dim(W) = \mathrm{rk}(\boldsymbol{A})$. A basis of W can be found by applying Gaussian elimination to $\boldsymbol{A}^\top$.

- For all $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ it holds that $\boldsymbol{A}$ is regular (invertible) if and only if $\text{rk}(\boldsymbol{A}) = n$.
- For all $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and all $\boldsymbol{b} \in \mathbb{R}^{m}$ it holds that the linear equation system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ can be solved if and only if $\text{rk}(\boldsymbol{A}) = \text{rk}(\boldsymbol{A}|\boldsymbol{b})$, where $\boldsymbol{A}|\boldsymbol{b}$ denotes the augmented system.
- For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ the subspace of solutions for $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ possesses dimension $n - \text{rk}(\boldsymbol{A})$. Later, we will call this subspace the *kernel* or the *null space*.
- A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ has *full rank* if its rank equals the largest possible rank for a matrix of the same dimensions. This means that the rank of a full-rank matrix is the lesser of the number of rows and columns, i.e., $\text{rk}(\boldsymbol{A}) = \min(m, n)$. A matrix is said to be *rank deficient* if it does not have full rank.

Example 2.18 (Rank)

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

  $\boldsymbol{A}$ has two linearly independent rows/columns so that $\text{rk}(\boldsymbol{A}) = 2$.

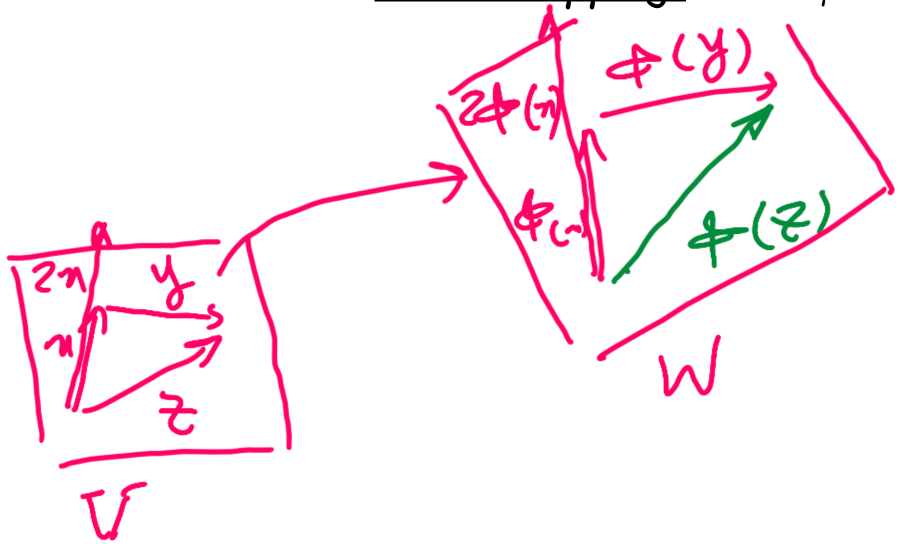
- $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix}$.

  We use Gaussian elimination to determine the rank:

$$\begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.84}$$

Here, we see that the number of linearly independent rows and columns is 2, such that $\text{rk}(\boldsymbol{A}) = 2$.

Linear Mappings



Consider two real vector spaces V, W. A mapping $\Phi : V \to W$ preserves the structure of the vector space if

$$\Phi(\boldsymbol{x}+\boldsymbol{y}) = \Phi(\boldsymbol{x}) + \Phi(\boldsymbol{y}) \tag{2.85}$$

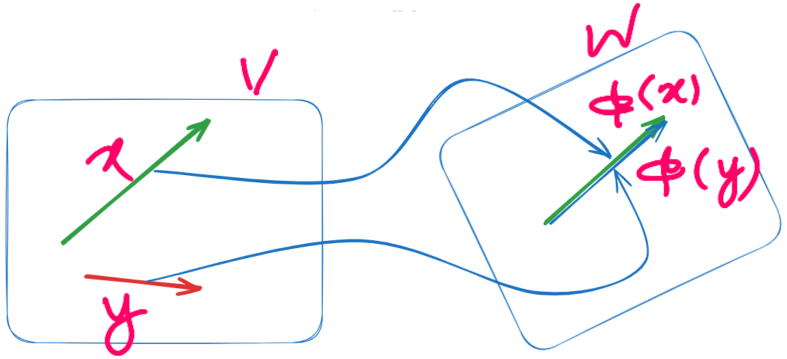
$$\Phi(\lambda\boldsymbol{x}) = \lambda\Phi(\boldsymbol{x}) \tag{2.86}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $\lambda \in \mathbb{R}$. We can summarize this in the following definition:

Definition 2.15 (Linear Mapping)**.** For vector spaces V, W, a mapping $\Phi : V \to W$ is called a *linear mapping* (or *vector space homomorphism/ linear transformation*) if

$$\forall \boldsymbol{x}, \boldsymbol{y} \in V \, \forall \lambda, \psi \in \mathbb{R} : \Phi(\lambda\boldsymbol{x} + \psi\boldsymbol{y}) = \lambda\Phi(\boldsymbol{x}) + \psi\Phi(\boldsymbol{y}) \,. \tag{2.87}$$

Definition 2.16 (Injective, Surjective, Bijective)**.** Consider a mapping $\Phi : \mathcal{V} \to \mathcal{W}$, where $\mathcal{V}, \mathcal{W}$ can be arbitrary sets. Then Φ is called



- *Injective* if $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V} : \Phi(\boldsymbol{x}) = \Phi(\boldsymbol{y}) \implies \boldsymbol{x} = \boldsymbol{y}$.
- *Surjective* if $\Phi(\mathcal{V}) = \mathcal{W}$.
- *Bijective* if it is injective and surjective.

using Φ. A bijective Φ can be "undone", i.e., there exists a mapping $\Psi : \mathcal{W} \to \mathcal{V}$ so that $\Psi \circ \Phi(\boldsymbol{x}) = \boldsymbol{x}$. This mapping Ψ is then called the inverse of Φ and normally denoted by Φ^{-1}.

همریختی

homomorphism

$\Phi(\lambda x+\psi y)=\lambda\Phi(x)+\psi\Phi(x)$

With these definitions, we introduce the following special cases of linear mappings between vector spaces V and W:

- *Isomorphism:* $\Phi : V \to W$ linear and bijective　همریختی دوسویه　یا همشکلی
- ~~*Endomorphism:* $\Phi : V \to V$~~ linear
- *Automorphism:* $\Phi : V \to V$ linear and bijective　خودریختی
- We define $\mathrm{id}_V : V \to V$, $\boldsymbol{x} \mapsto \boldsymbol{x}$ as the *identity mapping* or *identity automorphism* in V.

مفهوم	تعریف	ویژگی‌ها
Homomorphism	نگاشتی که ساختارهای جبری را حفظ می‌کند.	لزومی ندارد یک‌به‌یک یا پوشا باشد.
Isomorphism	همریختی که هم‌زمان یک‌به‌یک و پوشا است.	دو ساختار جبری را از نظر ساختاری یکسان می‌کند.
Automorphism	ایزومورفیسمی که دامنه و هم‌دامنه‌ی آن یکسان است (یعنی از یک ساختار به خودش).	ساختارهای جبری را حفظ می‌کند و یک‌به‌یک و پوشا است. مجموعه‌ی آن‌ها یک گروه است.

Example 2.19 (Homomorphism)
The mapping $\Phi : \mathbb{R}^2 \to \mathbb{C},\ \Phi(\boldsymbol{x}) = x_1 + ix_2$, is a homomorphism:

$$\begin{aligned} \Phi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) &= (x_1 + y_1) + i(x_2 + y_2) = x_1 + ix_2 + y_1 + iy_2 \\ &= \Phi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) + \Phi\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) \\ \Phi\left(\lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= \lambda x_1 + \lambda i x_2 = \lambda(x_1 + ix_2) = \lambda \Phi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) . \end{aligned} \tag{2.88}$$

This also justifies why complex numbers can be represented as tuples in $\mathbb{R}^2$: There is a bijective linear mapping that converts the elementwise addition of tuples in $\mathbb{R}^2$ into the set of complex numbers with the corresponding addition. Note that we only showed linearity, but not the bijection.

Theorem 2.17 (Theorem 3.59 in Axler (2015)). *Finite-dimensional vector spaces V and W are isomorphic if and only if* $\dim(V) = \dim(W)$.

Theorem 2.17 states that there exists a linear, bijective mapping between two vector spaces of the same dimension. Intuitively, this means that vector spaces of the same dimension are kind of the same thing, as they can be transformed into each other without incurring any loss.

Theorem 2.17 also gives us the justification to treat $\mathbb{R}^{m\times n}$ (the vector space of $m \times n$-matrices) and $\mathbb{R}^{mn}$ (the vector space of vectors of length mn) the same, as their dimensions are mn, and there exists a linear, bijective mapping that transforms one into the other.

Remark. Consider vector spaces V, W, X. Then:

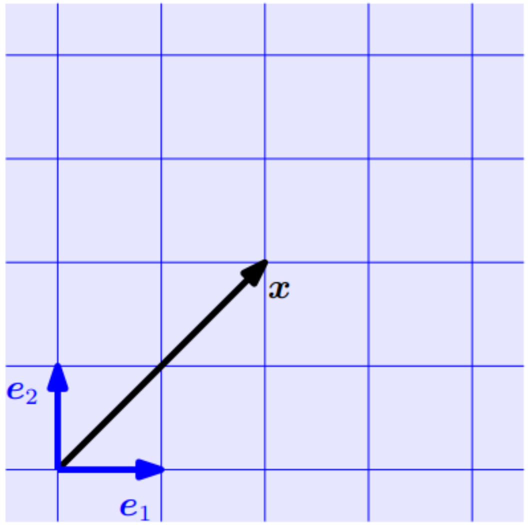
- For linear mappings $\Phi : V \to W$ and $\Psi : W \to X$, the mapping $\Psi \circ \Phi : V \to X$ is also linear.
- If $\Phi : V \to W$ is an isomorphism, then $\Phi^{-1} : W \to V$ is an isomorphism, too.
- If $\Phi : V \to W,\ \Psi : V \to W$ are linear, then $\Phi + \Psi$ and $\lambda\Phi,\ \lambda \in \mathbb{R}$, are linear, too.

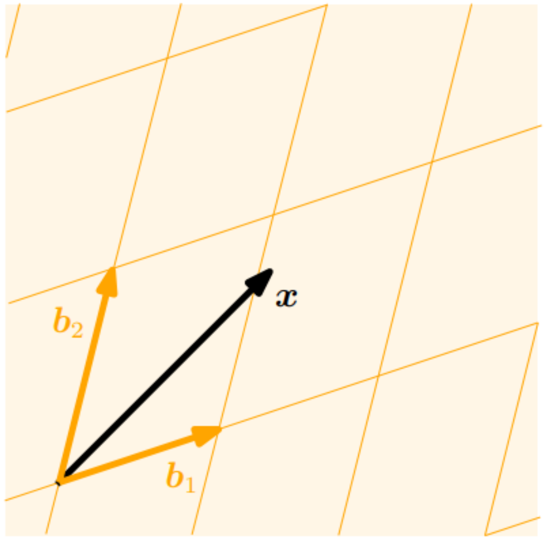
Matrix Representation of Linear Mappings

Any n-dimensional vector space is isomorphic to $\mathbb{R}^n$ (Theorem 2.17). We consider a basis $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_n\}$ of an n-dimensional vector space V. In the following, the order of the basis vectors will be important. Therefore, we write

$$B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n) \tag{2.89}$$

and call this n-tuple an *ordered basis* of V.





1. پایه مرتب: $B = (b_1, \dots, b_n)$
2. پایه نامرتب: $B = \{b_1, \dots, b_n\}$
3. یک ماتریس که ستون‌های آن بردارهای $b_1, \dots, b_n$ هستند: $B = [b_1, \dots, b_n]$

Definition 2.18 (Coordinates). Consider a vector space V and an ordered basis $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ of V. For any $\boldsymbol{x} \in V$ we obtain a unique representation (linear combination)

$$\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \ldots + \alpha_n \boldsymbol{b}_n \tag{2.90}$$

of $\boldsymbol{x}$ with respect to B. Then $\alpha_1, \ldots, \alpha_n$ are the *coordinates* of $\boldsymbol{x}$ with respect to B, and the vector

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n \tag{2.91}$$

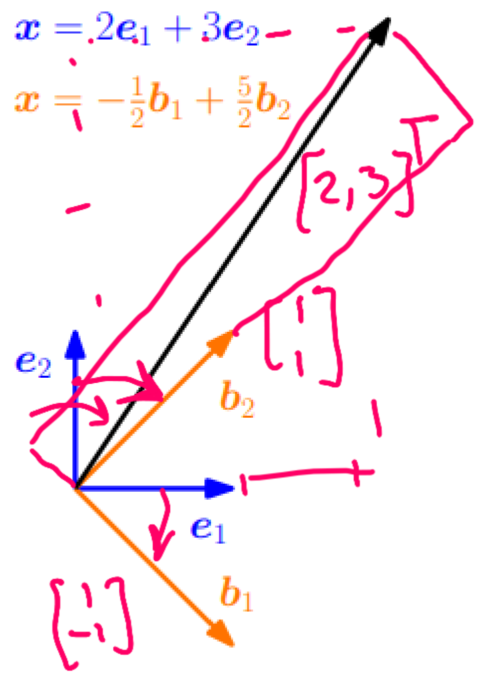
is the *coordinate vector/coordinate representation* of $\boldsymbol{x}$ with respect to the ordered basis B.

Example 2.20

Let us have a look at a geometric vector $\boldsymbol{x} \in \mathbb{R}^2$ with coordinates $[2, 3]^\top$ with respect to the standard basis $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ of $\mathbb{R}^2$. This means, we can write $\boldsymbol{x} = 2\boldsymbol{e}_1 + 3\boldsymbol{e}_2$. However, we do not have to choose the standard basis to represent this vector. If we use the basis vectors $\boldsymbol{b}_1 = [1, -1]^\top, \boldsymbol{b}_2 = [1, 1]^\top$ we will obtain the coordinates $\frac{1}{2}[-1, 5]^\top$ to represent the same vector with respect to $(\boldsymbol{b}_1, \boldsymbol{b}_2)$ (see Figure 2.9).

Remark. For an n-dimensional vector space V and an ordered basis B of V, the mapping $\Phi : \mathbb{R}^n \to V$, $\Phi(\boldsymbol{e}_i) = \boldsymbol{b}_i$, $i = 1, \ldots, n$, is linear (and because of Theorem 2.17 an isomorphism), where $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ is the standard basis of $\mathbb{R}^n$.

Different coordinate representations of a vector $\boldsymbol{x}$, depending on the choice of basis.



Definition 2.19 (Transformation Matrix). Consider vector spaces V, W with corresponding (ordered) bases $B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ and $C = (\boldsymbol{c}_1, \dots, \boldsymbol{c}_m)$. Moreover, we consider a linear mapping $\Phi : V \to W$. For $j \in \{1, \dots, n\}$,

$$\Phi(\boldsymbol{b}_j) = \alpha_{1j}\boldsymbol{c}_1 + \cdots + \alpha_{mj}\boldsymbol{c}_m = \sum_{i=1}^{m} \alpha_{ij}\boldsymbol{c}_i \tag{2.92}$$

is the unique representation of $\Phi(\boldsymbol{b}_j)$ with respect to C. Then, we call the $m \times n$-matrix $\boldsymbol{A}_\Phi$, whose elements are given by

$$A_\Phi(i, j) = \alpha_{ij} , \tag{2.93}$$

$[A_\Phi]_{n \times m}$

the *transformation matrix* of Φ (with respect to the ordered bases B of V and C of W).

$\vec{x} = \sum_{j=1}^{n} \hat{x}_j \vec{b}_j$

$\vec{y} = \sum_{i=1}^{m} \hat{y}_i \cancel{\vec{c}_i} = \Phi(\vec{x})$

$= \Phi\left(\sum_{j=1}^{n} \hat{x}_j \vec{b}_j\right) = \sum_{j=1}^{n} \hat{x}_j \Phi(\vec{b}_j)$

$= \sum_{j=1}^{n} \hat{x}_j \sum_{i=1}^{m} \alpha_{ij} \cancel{\vec{c}_i} \Rightarrow \sum_{i=1}^{m} \hat{y}_i = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} \hat{x}_j$

$\hat{y} = \alpha_{ij} \hat{x}_j$

transformation matrix $\boldsymbol{A}_\Phi$. If $\hat{\boldsymbol{x}}$ is the coordinate vector of $\boldsymbol{x} \in V$ with respect to B and $\hat{\boldsymbol{y}}$ the coordinate vector of $\boldsymbol{y} = \Phi(\boldsymbol{x}) \in W$ with respect to C, then

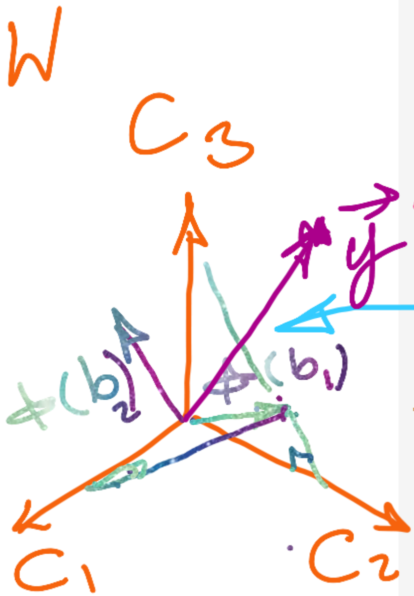
$$\boxed{\hat{\boldsymbol{y}} = \boldsymbol{A}_\Phi \hat{\boldsymbol{x}} .} \tag{2.94}$$

This means that the transformation matrix can be used to map coordinates with respect to an ordered basis in V to coordinates with respect to an ordered basis in W.

Example 2.21 (Transformation Matrix)

Consider a homomorphism $\Phi : V \to W$ and ordered bases $B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_3)$ of V and $C = (\boldsymbol{c}_1, \dots, \boldsymbol{c}_4)$ of W. With

$$A_\Phi = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}$$

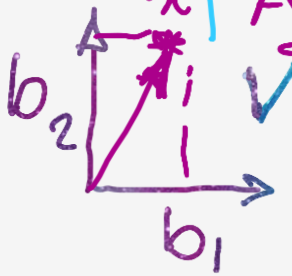


$$A_\Phi = [\alpha_{ij}]$$

$$\Phi$$

$$\begin{aligned} \Phi(\boldsymbol{b}_1) &= \boldsymbol{c}_1 - \boldsymbol{c}_2 + 3\boldsymbol{c}_3 - \boldsymbol{c}_4 \\ \Phi(\boldsymbol{b}_2) &= 2\boldsymbol{c}_1 + \boldsymbol{c}_2 + 7\boldsymbol{c}_3 + 2\boldsymbol{c}_4 \\ \Phi(\boldsymbol{b}_3) &= 3\boldsymbol{c}_2 + \boldsymbol{c}_3 + 4\boldsymbol{c}_4 \end{aligned} \tag{2.95}$$

the transformation matrix $\boldsymbol{A}_\Phi$ with respect to B and C satisfies $\Phi(\boldsymbol{b}_k) = \sum_{i=1}^{4} \alpha_{ik} \boldsymbol{c}_i$ for $k = 1, \dots, 3$ and is given as

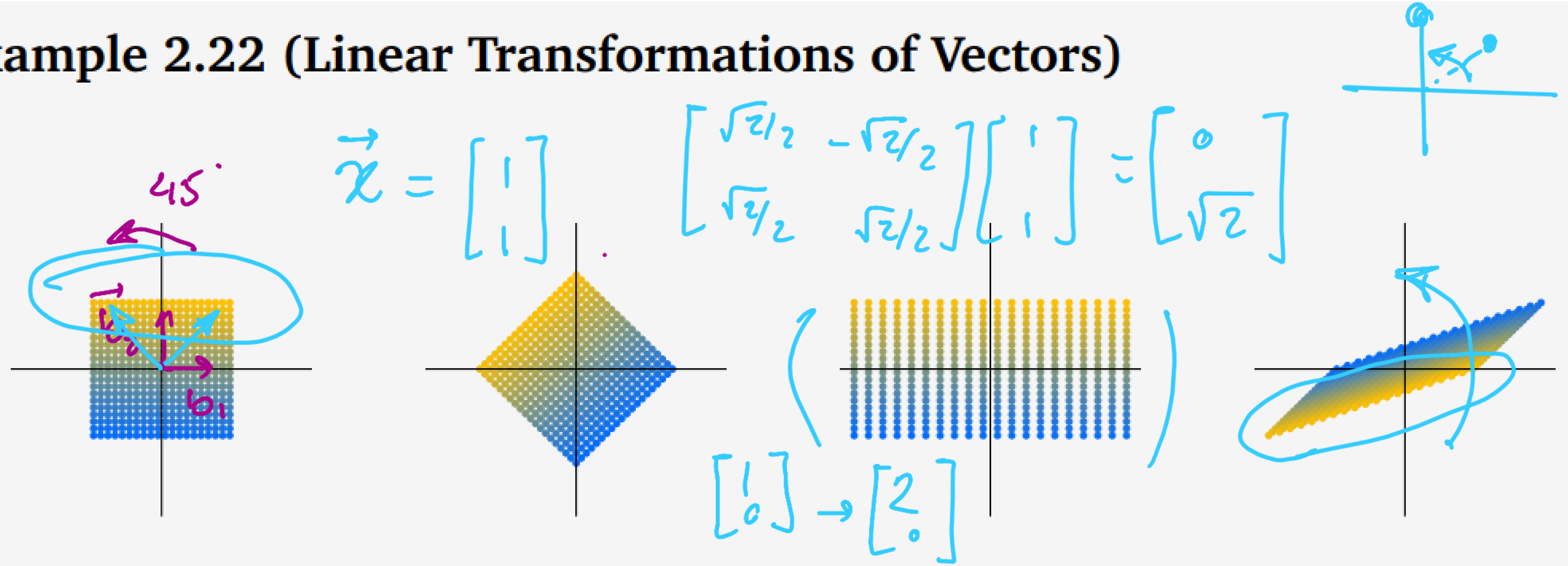


$$A_\Phi \hat{x} = \hat{y}$$

$$\boldsymbol{A}_\Phi = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3] = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}, \tag{2.96}$$

where the $\boldsymbol{\alpha}_j$, $j = 1, 2, 3$, are the coordinate vectors of $\Phi(\boldsymbol{b}_j)$ with respect to C.

Example 2.22 (Linear Transformations of Vectors)



(a) Original data. (b) Rotation by 45°. (c) Stretch along the horizontal axis. (d) General linear mapping.

We consider three linear transformations of a set of vectors in $\mathbb{R}^2$ with the transformation matrices

$$A_1 = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_3 = \frac{1}{2}\begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix}. \tag{2.97}$$

Basis Change

Theorem 2.20 (Basis Change). *For a linear mapping $\Phi : V \to W$, ordered bases*

$$B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n), \quad \tilde{B} = (\tilde{\boldsymbol{b}}_1, \ldots, \tilde{\boldsymbol{b}}_n) \tag{2.103}$$

of V and

$$C = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m), \quad \tilde{C} = (\tilde{\boldsymbol{c}}_1, \ldots, \tilde{\boldsymbol{c}}_m) \tag{2.104}$$

of W, and a transformation matrix $\boldsymbol{A}_\Phi$ of Φ with respect to B and C, the corresponding transformation matrix $\tilde{\boldsymbol{A}}_\Phi$ with respect to the bases $\tilde{B}$ and $\tilde{C}$ is given as

$$\boxed{\tilde{\boldsymbol{A}}_\Phi = \boldsymbol{T}^{-1}\boldsymbol{A}_\Phi\boldsymbol{S}\,.} \quad \underset{m\times m}{[T^{-1}]}\,\underset{m\times n}{[A_\Phi]}\,\underset{n\times n}{[S]} \tag{2.105}$$

Here, $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ is the transformation matrix of id_V that maps coordinates with respect to $\tilde{B}$ onto coordinates with respect to B, and $\boldsymbol{T} \in \mathbb{R}^{m\times m}$ is the transformation matrix of id_W that maps coordinates with respect to $\tilde{C}$ onto coordinates with respect to C.

a.fahim@ut.ac.ir

Basis Change

Vector spaces

Ordered bases

$$\begin{tikzcd}
V \arrow[rr, "\Phi"] & & W \\
B \arrow[rr, "\Phi_{CB}", "A_{\Phi}"'] & & C \\
\tilde{B} \arrow[u, "\Psi_{B\tilde{B}}", "S"'] \arrow[rr, "\tilde{A}_{\Phi}", "\Phi_{\tilde{C}\tilde{B}}"'] & & \tilde{C} \arrow[u, "T", "\Xi_{C\tilde{C}}"']
\end{tikzcd}
\qquad
\begin{tikzcd}
V \arrow[rr, "\Phi"] & & W \\
B \arrow[rr, "\Phi_{CB}", "A_{\Phi}"'] & & C \arrow[d, "T^{-1}"', "\Xi_{\tilde{C}C} = \Xi^{-1}_{C\tilde{C}}"] \\
\tilde{B} \arrow[u, "\Psi_{B\tilde{B}}", "S"'] \arrow[rr, "\tilde{A}_{\Phi}", "\Phi_{\tilde{C}\tilde{B}}"'] & & \tilde{C}
\end{tikzcd}$$

Figure 2.11 For a homomorphism $\Phi : V \to W$ and ordered bases $B, \tilde{B}$ of V and $C, \tilde{C}$ of W (marked in blue), we can express the mapping $\Phi_{\tilde{C}\tilde{B}}$ with respect to the bases $\tilde{B}, \tilde{C}$ equivalently as a composition of the homomorphisms $\Phi_{\tilde{C}\tilde{B}} = \Xi_{\tilde{C}C} \circ \Phi_{CB} \circ \Psi_{B\tilde{B}}$ with respect to the bases in the subscripts. The corresponding transformation matrices are in red.

Therefore, we can express the linear mapping $\Phi_{\tilde{C}\tilde{B}}$ as a composition of linear mappings that involve the "old" basis:

$$\Phi_{\tilde{C}\tilde{B}} = \Xi_{\tilde{C}C} \circ \Phi_{CB} \circ \Psi_{B\tilde{B}} = \Xi^{-1}_{C\tilde{C}} \circ \Phi_{CB} \circ \Psi_{B\tilde{B}} \,. \tag{2.114}$$

Concretely, we use $\Psi_{B\tilde{B}} = \mathrm{id}_V$ and $\Xi_{C\tilde{C}} = \mathrm{id}_W$, i.e., the identity mappings that map vectors onto themselves, but with respect to a different basis.

Basis Change

Definition 2.21 (Equivalence). Two matrices $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{m\times n}$ are *equivalent* if there exist regular matrices $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{T} \in \mathbb{R}^{m\times m}$, such that $\underbrace{\tilde{\boldsymbol{A}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{S}}$.

Definition 2.22 (Similarity). Two matrices $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{n\times n}$ are *similar* if there exists a regular matrix $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ with $\tilde{\boldsymbol{A}} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$

Remark. Similar matrices are always equivalent. However, equivalent matrices are not necessarily similar. ◊

Remark. Consider vector spaces V, W, X. From the remark that follows Theorem 2.17, we already know that for linear mappings $\Phi : V \to W$ and $\Psi : W \to X$ the mapping $\Psi \circ \Phi : V \to X$ is also linear. With transformation matrices $\boldsymbol{A}_\Phi$ and $\boldsymbol{A}_\Psi$ of the corresponding mappings, the overall transformation matrix is $\underbrace{\boldsymbol{A}_{\Psi\circ\Phi}} = \underbrace{\boldsymbol{A}_\Psi \boldsymbol{A}_\Phi}$. ◊

a.fahim@ut.ac.ir

Basis Change

- $\boldsymbol{A}_\Phi$ is the transformation matrix of a linear mapping $\Phi_{CB} : V \to W$ with respect to the bases B, C.
- $\tilde{\boldsymbol{A}}_\Phi$ is the transformation matrix of the linear mapping $\Phi_{\tilde{C}\tilde{B}} : V \to W$ with respect to the bases $\tilde{B}, \tilde{C}$.
- $\boldsymbol{S}$ is the transformation matrix of a linear mapping $\Psi_{B\tilde{B}} : V \to V$ (automorphism) that represents $\tilde{B}$ in terms of B. Normally, $\Psi = \mathrm{id}_V$ is the identity mapping in V.
- $\boldsymbol{T}$ is the transformation matrix of a linear mapping $\Xi_{C\tilde{C}} : W \to W$ (automorphism) that represents $\tilde{C}$ in terms of C. Normally, $\Xi = \mathrm{id}_W$ is the identity mapping in W.

Basis Change

If we (informally) write down the transformations just in terms of bases, then $\boldsymbol{A}_\Phi : B \to C$, $\tilde{\boldsymbol{A}}_\Phi : \underbrace{\tilde{B} \to \tilde{C}}$, $\boldsymbol{S} : \tilde{B} \to B$, $\boldsymbol{T} : \tilde{C} \to C$ and $\boldsymbol{T}^{-1} : C \to \tilde{C}$, and

$$\tilde{B} \to \tilde{C} = \tilde{B} \to B \to C \to \tilde{C} \tag{2.115}$$

$$\tilde{\boldsymbol{A}}_\Phi = \boldsymbol{T}^{-1} \boldsymbol{A}_\Phi \boldsymbol{S} \,. \tag{2.116}$$

Note that the execution order in (2.116) is from right to left because vectors are multiplied at the right-hand side so that $\boldsymbol{x} \mapsto \boldsymbol{S}\boldsymbol{x} \mapsto \boldsymbol{A}_\Phi(\boldsymbol{S}\boldsymbol{x}) \mapsto \boldsymbol{T}^{-1}(\boldsymbol{A}_\Phi(\boldsymbol{S}\boldsymbol{x})) = \tilde{\boldsymbol{A}}_\Phi \boldsymbol{x}$.

V $\qquad$ W

$\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ $\qquad$ $\tilde{C} = (\tilde{c}_1, \ldots, \tilde{c}_m)$

a.fahim@ut.ac.ir

Basis Change

Example 2.24 (Basis Change)
Consider a linear mapping $\Phi : \mathbb{R}^3 \to \mathbb{R}^4$ whose transformation matrix is

$$A_\Phi = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix} \quad (2.117)$$

4x3

with respect to the standard bases

$$B = \left(\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right), \quad C = \left(\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right). \quad (2.118)$$

We seek the transformation matrix $\tilde{A}_\Phi$ of Φ with respect to the new bases

S

$$\tilde{B} = \left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right) \in \mathbb{R}^3, \quad \tilde{C} = \left(\begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\1\end{bmatrix}\right). \quad (2.119)$$

Then,

$[T \mid I_4] \rightsquigarrow [I_4 \mid T^{-1}]$

$$S = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (2.120)$$

where the ith column of S is the coordinate representation of $\tilde{b}_i$ in terms of the basis vectors of B. Since B is the standard basis, the coordinate representation is straightforward to find. For a general basis B, we would need to solve a linear equation system to find the λ_i such that

$\sum_{i=1}^{3} \lambda_i b_i = \tilde{b}_j$, $j = 1, \ldots, 3$. Similarly, the jth column of T is the coordinate representation of $\tilde{c}_j$ in terms of the basis vectors of C.

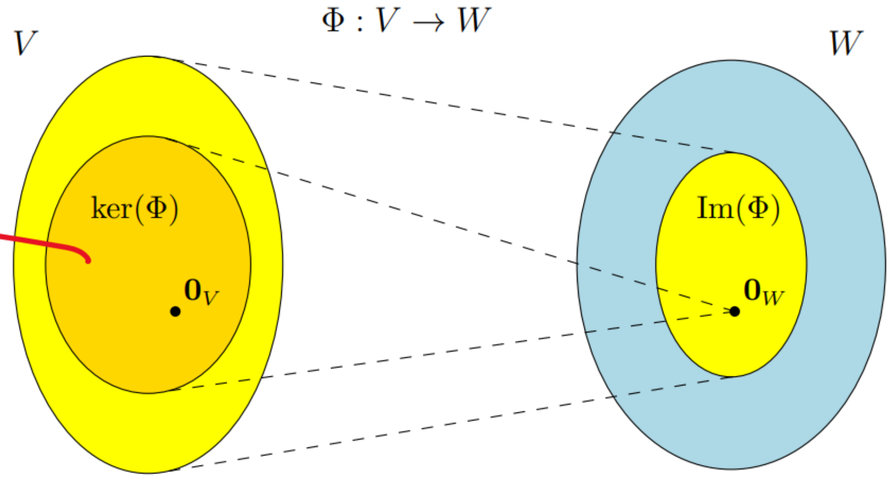
Therefore, we obtain

$$\underline{\underline{\tilde{A}_\Phi = T^{-1} A_\Phi S}} = \frac{1}{2} \underbrace{\begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}}_{T^{-1}} \underbrace{\begin{bmatrix} 3 & 2 & 1 \\ 0 & 4 & 2 \\ 10 & 8 & 4 \\ 1 & 6 & 3 \end{bmatrix}}_{A_\Phi S} \quad (2.121a)$$

$$= \begin{bmatrix} -4 & -4 & -2 \\ 6 & 0 & 0 \\ 4 & 8 & 4 \\ 1 & 6 & 3 \end{bmatrix}. \quad (2.121b)$$

Image and Kernel

null space



Definition 2.23 (Image and Kernel).
For $\Phi : V \to W$, we define the *kernel/null space*

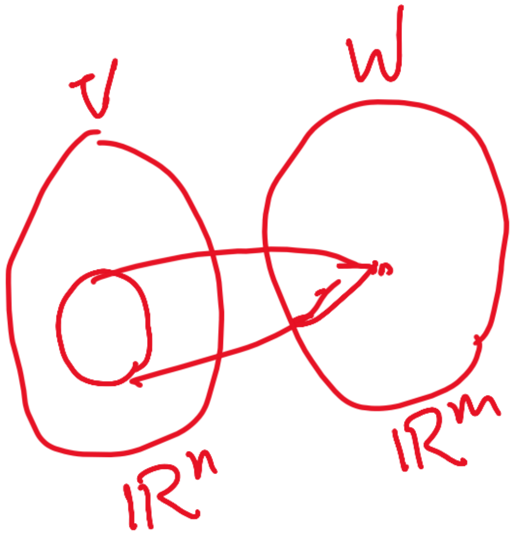
$$\ker(\Phi) := \Phi^{-1}(\mathbf{0}_W) = \{\boldsymbol{v} \in V : \Phi(\boldsymbol{v}) = \mathbf{0}_W\} \tag{2.122}$$

and the *image/range*

$$\mathrm{Im}(\Phi) := \Phi(V) = \{\boldsymbol{w} \in W | \exists \boldsymbol{v} \in V : \Phi(\boldsymbol{v}) = \boldsymbol{w}\} . \tag{2.123}$$

We also call V and W also the *domain* and *codomain* of Φ, respectively.

- It always holds that $\Phi(\mathbf{0}_V) = \mathbf{0}_W$ and, therefore, $\mathbf{0}_V \in \ker(\Phi)$. In particular, the null space is never empty.
- $\mathrm{Im}(\Phi) \subseteq W$ is a subspace of W, and $\ker(\Phi) \subseteq V$ is a subspace of V.
- Φ is injective (one-to-one) if and only if $\ker(\Phi) = \{\mathbf{0}\}$.



$A_{m\times n}$

Remark (Null Space and Column Space). Let us consider $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and a linear mapping $\Phi : \overset{V}{\mathbb{R}^n} \to \overset{W}{\mathbb{R}^m}$, $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$.

- For $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]$, where $\boldsymbol{a}_i$ are the columns of $\boldsymbol{A}$, we obtain

$$\mathrm{Im}(\Phi) = \{\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n\} = \left\{\sum_{i=1}^{n} x_i \boldsymbol{a}_i : x_1, \ldots, x_n \in \mathbb{R}\right\} \tag{2.124a}$$

$$= \mathrm{span}[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n] \subseteq \mathbb{R}^m\,, \tag{2.124b}$$

  i.e., the image is the span of the columns of $\boldsymbol{A}$, also called the *column space*. Therefore, the column space (image) is a subspace of $\mathbb{R}^m$, where m is the “height” of the matrix.
- $\mathrm{rk}(\boldsymbol{A}) = \dim(\mathrm{Im}(\Phi))$.
- The kernel/null space $\ker(\Phi)$ is the general solution to the homogeneous system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ and captures all possible linear combinations of the elements in $\mathbb{R}^n$ that produce $\boldsymbol{0} \in \mathbb{R}^m$.
- The kernel is a subspace of $\mathbb{R}^n$, where n is the “width” of the matrix.
- The kernel focuses on the relationship among the columns, and we can use it to determine whether/how we can express a column as a linear combination of other columns.

Example 2.25 (Image and Kernel of a Linear Mapping)
The mapping

$$\Phi : \mathbb{R}^4 \to \mathbb{R}^2, \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mapsto \underbrace{\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\vec{x} \in V} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix} \quad W \tag{2.125a}$$

$$\text{Ker}(\Phi) = \text{Span}\left[\begin{bmatrix} 0 \\ -1/2 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1/2 \\ 0 \\ -1 \end{bmatrix}\right]$$

$$= x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.125b}$$

is linear. To determine $\text{Im}(\Phi)$, we can take the span of the columns of the transformation matrix and obtain

$$\text{Im}(\Phi) = \text{span}\left[\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right]. \quad W \tag{2.126}$$

To compute the kernel (null space) of Φ, we need to solve $\boldsymbol{Ax} = \boldsymbol{0}$, i.e., we need to solve a homogeneous equation system. To do this, we use Gaussian elimination to transform $\boldsymbol{A}$ into reduced row-echelon form:

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}. \tag{2.127}$$

This matrix is in reduced row-echelon form, and we can use the Minus-1 Trick to compute a basis of the kernel (see Section 2.3.3). Alternatively, we can express the non-pivot columns (columns 3 and 4) as linear combinations of the pivot columns (columns 1 and 2). The third column $\boldsymbol{a}_3$ is equivalent to $-\frac{1}{2}$ times the second column $\boldsymbol{a}_2$. Therefore, $\boldsymbol{0} = \boldsymbol{a}_3 + \frac{1}{2}\boldsymbol{a}_2$. In the same way, we see that $\boldsymbol{a}_4 = \boldsymbol{a}_1 - \frac{1}{2}\boldsymbol{a}_2$ and, therefore, $\boldsymbol{0} = \boldsymbol{a}_1 - \frac{1}{2}\boldsymbol{a}_2 - \boldsymbol{a}_4$. Overall, this gives us the kernel (null space) as

$$\ker(\Phi) = \text{span}\left[\begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}\right]. \tag{2.128}$$

Theorem 2.24 (Rank-Nullity Theorem). *For vector spaces V, W and a linear mapping $\Phi : V \to W$ it holds that*

$$\dim(\ker(\Phi)) + \dim(\mathrm{Im}(\Phi)) = \dim(V)\,. \tag{2.129}$$

The rank-nullity theorem is also referred to as the *fundamental theorem of linear mappings* (Axler, 2015, theorem 3.22). The following are direct consequences of Theorem 2.24:

- If $\dim(\mathrm{Im}(\Phi)) < \dim(V)$, then $\ker(\Phi)$ is non-trivial, i.e., the kernel contains more than $\mathbf{0}_V$ and $\dim(\ker(\Phi)) \geqslant 1$.
- If $\boldsymbol{A}_\Phi$ is the transformation matrix of Φ with respect to an ordered basis and $\dim(\mathrm{Im}(\Phi)) < \dim(V)$, then the system of linear equations $\boldsymbol{A}_\Phi \boldsymbol{x} = \mathbf{0}$ has infinitely many solutions.
- If $\dim(V) = \dim(W)$, then the three-way equivalence

$$\Phi \text{ is injective} \iff \Phi \text{ is surjective} \iff \Phi \text{ is bijective}$$

holds since $\mathrm{Im}(\Phi) \subseteq W$.

$$\boldsymbol{A}_\Phi = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}$$

4×3 → n

$\Phi : \mathbb{R}^3 \to \mathbb{R}^4$

$$A_\Phi = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\mathrm{rk}(A_\Phi) = 2 \Rightarrow \dim(\mathrm{Im}(\Phi)) = 2$

$\dim(\ker(\Phi)) = n - \dim(\mathrm{Im}(\Phi)) = 3 - 2 = 1$

Definition 2.25 (Affine Subspace). Let V be a vector space, $\boldsymbol{x}_0 \in V$ and $U \subseteq V$ a subspace. Then the subset

$$L = \boldsymbol{x}_0 + U := \{\boldsymbol{x}_0 + \boldsymbol{u} : \boldsymbol{u} \in U\} \tag{2.130a}$$

$$= \{\boldsymbol{v} \in V | \exists \boldsymbol{u} \in U : \boldsymbol{v} = \boldsymbol{x}_0 + \boldsymbol{u}\} \subseteq V \tag{2.130b}$$

is called *affine subspace* or *linear manifold* of V. U is called *direction* or *direction space*, and $\boldsymbol{x}_0$ is called *support point*.

Note that the definition of an affine subspace excludes $\mathbf{0}$ if $\boldsymbol{x}_0 \notin U$. Therefore, an affine subspace is not a (linear) subspace (vector subspace) of V for $\boldsymbol{x}_0 \notin U$.

Examples of affine subspaces are points, lines, and planes in $\mathbb{R}^3$, which do not (necessarily) go through the origin.

Affine subspaces are often described by *parameters*: Consider a k-dimensional affine space $L = \boldsymbol{x}_0 + U$ of V. If $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k)$ is an ordered basis of U, then every element $\boldsymbol{x} \in L$ can be uniquely described as

$$\boldsymbol{x} = \boldsymbol{x}_0 + \lambda_1 \boldsymbol{b}_1 + \ldots + \lambda_k \boldsymbol{b}_k \,, \tag{2.131}$$

where $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. This representation is called *parametric equation* of L with directional vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ and *parameters* $\lambda_1, \ldots, \lambda_k$. ◇

Example 2.26 (Affine Subspaces)

- One-dimensional affine subspaces are called *lines* and can be written as $\boldsymbol{y} = \boldsymbol{x}_0 + \lambda \boldsymbol{b}_1$, where $\lambda \in \mathbb{R}$ and $U = \text{span}[\boldsymbol{b}_1] \subseteq \mathbb{R}^n$ is a one-dimensional subspace of $\mathbb{R}^n$. This means that a line is defined by a support point $\boldsymbol{x}_0$ and a vector $\boldsymbol{b}_1$ that defines the direction. See Figure 2.13 for an illustration.

- Two-dimensional affine subspaces of $\mathbb{R}^n$ are called *planes*. The parametric equation for planes is $\boldsymbol{y} = \boldsymbol{x}_0 + \lambda_1 \boldsymbol{b}_1 + \lambda_2 \boldsymbol{b}_2$, where $\lambda_1, \lambda_2 \in \mathbb{R}$ and $U = \text{span}[\boldsymbol{b}_1, \boldsymbol{b}_2] \subseteq \mathbb{R}^n$. This means that a plane is defined by a support point $\boldsymbol{x}_0$ and two linearly independent vectors $\boldsymbol{b}_1, \boldsymbol{b}_2$ that span the direction space.
- In $\mathbb{R}^n$, the $(n-1)$-dimensional affine subspaces are called *hyperplanes*, and the corresponding parametric equation is $\boldsymbol{y} = \boldsymbol{x}_0 + \sum_{i=1}^{n-1} \lambda_i \boldsymbol{b}_i$, where $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-1}$ form a basis of an $(n-1)$-dimensional subspace U of $\mathbb{R}^n$. This means that a hyperplane is defined by a support point $\boldsymbol{x}_0$ and $(n-1)$ linearly independent vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-1}$ that span the direction space. In $\mathbb{R}^2$, a line is also a hyperplane. In $\mathbb{R}^3$, a plane is also a hyperplane.

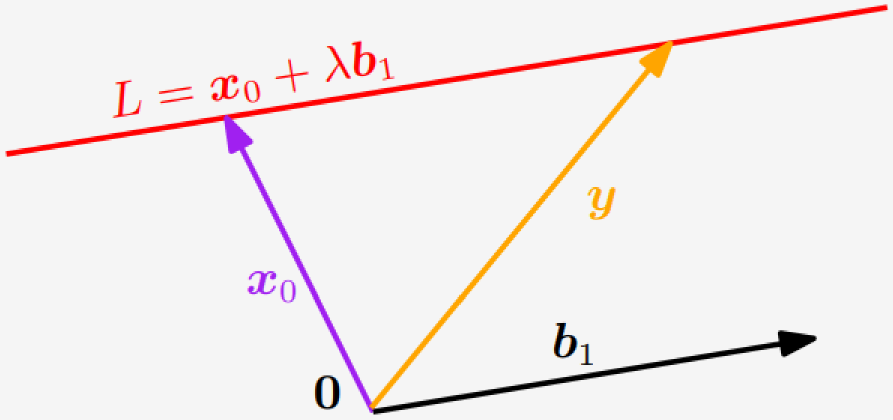


Figure 2.13 Lines are affine subspaces. Vectors $\boldsymbol{y}$ on a line $\boldsymbol{x}_0 + \lambda \boldsymbol{b}_1$ lie in an affine subspace L with support point $\boldsymbol{x}_0$ and direction $\boldsymbol{b}_1$.

Remark (Inhomogeneous systems of linear equations and affine subspaces). For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{x} \in \mathbb{R}^m$, the solution of the system of linear equations $\boldsymbol{A\lambda} = \boldsymbol{x}$ is either the empty set or an affine subspace of $\mathbb{R}^n$ of dimension $n - \text{rk}(\boldsymbol{A})$. In particular, the solution of the linear equation $\lambda_1 \boldsymbol{b}_1 + \ldots + \lambda_n \boldsymbol{b}_n = \boldsymbol{x}$, where $(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0)$, is a hyperplane in $\mathbb{R}^n$.

In $\mathbb{R}^n$, every k-dimensional affine subspace is the solution of an inhomogeneous system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$, where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ and $\text{rk}(\boldsymbol{A}) = n - k$. Recall that for homogeneous equation systems $\boldsymbol{Ax} = \boldsymbol{0}$ the solution was a vector subspace, which we can also think of as a special affine space with support point $\boldsymbol{x}_0 = \boldsymbol{0}$. $\diamondsuit$

Definition 2.26 (Affine Mapping)**.** For two vector spaces V, W, a linear mapping $\Phi : V \to W$, and $\boldsymbol{a} \in W$, the mapping

$$\phi : V \to W \tag{2.132}$$

$$\boldsymbol{x} \mapsto \boldsymbol{a} + \Phi(\boldsymbol{x}) \tag{2.133}$$

is an *affine mapping* from V to W. The vector $\boldsymbol{a}$ is called the *translation vector* of ϕ.

- Every affine mapping $\phi : V \to W$ is also the composition of a linear mapping $\Phi : V \to W$ and a translation $\tau : W \to W$ in W, such that $\phi = \tau \circ \Phi$. The mappings Φ and τ are uniquely determined.
- The composition $\phi' \circ \phi$ of affine mappings $\phi : V \to W$, $\phi' : W \to X$ is affine.
- If ϕ is bijective, affine mappings keep the geometric structure invariant. They then also preserve the dimension and parallelism.