Contents

Chapter 1

Linear Systems

1.1 Linearity and linear equations

Definition 1 (linear combination)**.** *A* linear combination *of $x_1 \ldots x_n$ is an expression of the form*

$$a_1x_1 + a_2x_2 + \cdots a_nx_n,$$

where x_i's are indeterminates *or* variables*; and a_i's are* coefficients *and belong to a field $\mathbb{K}$.*

Remark: In the above definition and in the remainder of these notes the field $\mathbb{K}$ will be mostly either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$, but in general any field will do, except in special cases when $\mathbb{C}$ is required. We will strive to make a note whenever $\mathbb{C}$ is required.

Examples:

linear	non-linear
$x + y + z$	$\tan(x) + x + (-1)z$
$2ix + y + (-1)z$	$x + yx + (-1)z$
$x_1 + (1 - 7i)y + \sqrt{-1}z$	$x^2 + \sin(y)x + (-1)z$
$\left(\int_0^4 x\,\mathrm{d}x\right) x_1 + (1 - 7i)x_2 + \sqrt{-1}x_3$	$\sin(x) + \sin(y) + \sin(z)$

Remark: often $0x$ is omitted. Likewise instead of $(-1)x$ one writes $-x$. There is a difference between linear in x and linear in $\sin x$.

Definition 2 (linear equation)**.** *A* linear equation *in the set of variables $\mathbf{X}$, where without loss of generality $\mathbf{X} = \{x_1 \ldots x_n\}$ is an equation of the form*

$$a_1x_1 + a_2x_2 + \cdots a_nx_n = b.$$

The value b is the constant *of the linear equation and similar to the coefficients it belongs to $\mathbb{K}$.*

Notation: when referring to the linear equation

$$a_1x_1 + a_2x_2 + \cdots + a_nc_n = b$$

we will use the summation symbol $\sum$ and write

$$\sum_{i=1}^{n} a_ix_i = b.$$

The limits of the summation can be any two numbers (negatives are fine for example). We sum only integers indices.

Examples: For a set of variable $\mathbf{X} = \{x_1, x_2, x_3, x_4\}$ we have

linear equations	*not* linear equations
$x_1 + 4x_2 + x_3 + 2x_4 = -8$	$\sqrt{x_1} + 3x_2 + x_4 = 3$
$2x_1 + 3x_2 + x_3 = 5$	$x_1x_2 + x_2 + x_4 = 6$

Definition 3 (homogeneous equation)**.** *An equation $\sum_{i=1}^{n} a_ix_i = b$ is called* homogeneous *if the constant of the equation i.e., b is zero, that is $b = 0$.*

Examples:

non-homogeneous linear equations	homogeneous linear equations
$x_1 + 4x_2 + x_3 + 2x_4 = -8$	$x_1 + 4x_2 + x_3 + 2x_4 = 0$
$2x_1 + 3x_2 + x_3 = 5$	$2x_1 + 3x_2 + x_3 = 0$

1.1.1 Solutions to a linear equation

Definition 4 (solution of an equation)**.** *An n-tuple $(s_1, \ldots, s_n) \in \mathbb{K}^n$ is a solution*[1] *to the linear equation*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \iff a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$$

If the set of solution is empty, then the equation is said to have no solution or is inconsistent. *If it has at least one solution it is said to be* consistent.

Observe that the second equation is concerned with numbers only; there are no indeterminates.

Examples: $(1, 0, -3)$ is solutions to $4x_1 + 2x_2 + x_3 = 1$, but $(-3, 0, 1)$ is not a solution.

Example: The values $(-2, 5, 0)$ and $(0, 4, -1)$ are both solutions to $x_1 + x_2 + x_3 = 3$, $(1, 5, 0)$ is *not* a solution to $x_1 + x_2 + x_3 = 3$. The set of all solutions is $x_3 = s_1, x_2 = s_2$ and $x_1 = 3 - s_1 - s_2$ where $s_1, s_2 \in \mathbb{K}$.

[1]Order is important

Example: the solution to $x_1 + x_2 + x_3 = 1$ is the set

$$\{(1 - s_1 - s_2, s_1, s_2) \mid s_1, s_2 \in \mathbb{K}\} \subseteq \mathbb{K}^3$$

Theorem 1. *Let l be the linear equation*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

1. *if at least one a_i is non-zero then the solution set is*

$$\begin{aligned} \mathcal{S} \quad = \quad & \left\{ \left(s_1, \ldots, s_{i-1}, \frac{b - (a_1s_1 + \cdots + a_{i-1}s_{i-1} + a_{i+1}s_{i+1} + \cdots + a_ns_n)}{a_i}, s_{i+1}, \ldots, s_n \right) \right. \\ & \qquad \left. \mid s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n \in \mathbb{K} \right\}; \end{aligned}$$

2. *if all a_i's are zero and*

   (a) *the constant b is zero then the solution set is*

   $$\{(s_1, \ldots, s_n,) \mid s_1, \ldots, s_n \in \mathbb{K}\}\,;$$

   (b) *and lastly if all a_i's are zero and the constant b is non-zero then the solution set is empty.*

Proof. Suppose $a_i \neq 0$ and let $s_1, \ldots, s_n$ be a solution to the linear equation, by definition

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$$

rearranging

$$s_i = \frac{b - (a_1s_1 + \cdots + a_{i-1}s_{i-1} + a_{i+1}s_{i+1} + \cdots + a_ns_n)}{a_i}.$$

Thus $s_1, \ldots, s_n \in \mathcal{S}$. Take an element from the set $\mathcal{S}$ and consider

$$\begin{aligned} & a_1s_1 + \cdots + a_{i-1}s_{i-1} \\ & \quad + a_i\frac{b - (a_1s_1 + \cdots + a_{i-1}s_{i-1} + a_{i+1}s_{i+1} + \cdots + a_ns_n)}{a_i} \\ & \quad\quad + a_{i+1}s_{i+1} + \cdots + a_ns_n \\ = \quad & a_1s_1 + \cdots + a_{i-1}s_{i-1} + a_{i+1}s_{i+1} + \cdots + a_ns_n \\ & \quad + b - (a_1s_1 + \cdots + a_{i-1}s_{i-1} + a_{i+1}s_{i+1} + \cdots + a_ns_n) \\ = \quad & b \end{aligned}$$

Thus every solution belongs to the set $\mathcal{S}$ and every element of the set $\mathcal{S}$ is a solution to the linear equation concluding this part of the argument.

Suppose now all a_i's are zero that is $a_1 = 0$, $a_2 = 0 \ldots a_n = 0$. For any n-tuple $(s_1, \ldots, s_n)$ we have

$$a_1 s_1 + \cdots + a_n s_n = 0s_1 + \cdots + 0s_n = 0(s_1 + \cdots + s_n) = 0$$

Therefore if $b = 0$ every n-tuple is a solution. If $b \neq 0$ the equation has no solution, concluding the argument.

$\square$

Example: the solution to $0x_1 + 2x_2 + x_3 = 3$ is the set

$$\left\{ \left(s_1, \frac{3 - s_2}{2}, s_2 \right) \mid s_1, s_2 \in \mathbb{K} \right\} \subseteq \mathbb{K}^3.$$

The description of the solution set is not unique. The same set – the set of solution for this equation can also be given as

$$\{(t_1, t_2, 3 - 2t_2) \mid t_1, t_2 \in \mathbb{K}\} \subseteq \mathbb{K}^3$$

Remark. Typically in case there is a non-zero coefficient the solution set is described using the smallest index i for which a_i is non-zero. For the above example it means the solution set is described with the former description instead of the latter.

Example: the solution to $0x_1 + 0x_2 = 0$ is $\{(s_1, s_2) \mid s_1, s_2 \in \mathbb{K}\} \subseteq \mathbb{K}^2$

Example: the solution to $0x_1 + 0x_2 + 0x_3 + 0x_4 = 2$ is the empty set $\emptyset \subseteq \mathbb{K}^4$ in other words it has no solution, in other words the equation is inconsistent.

Example: the solution to $2x_1 = 6$ is the set with single element $3 \in \mathbb{K}^1$. It is sometimes called singleton, the set can be denoted by $\{3\} \subseteq \mathbb{K}^1$.

1.2 System of linear equations

Definition 5 (system of linear equations)**.** *A system of linear equation is a set of linear equations in the same set of variables* $\mathbf{X} = \{x_1, \ldots, x_n\}$*:*

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

Definition 6 (solution to a system of linear equations)**.** *An n-tuple* $(s_1, \ldots, s_n) \in \mathbb{K}^n$ *is a solution to a linear system of equations if it is solution for each equation.*

Remark: The above definition implies that the set of solutions to a system of linear equations is the intersection of the sets of solutions for each equation in the system of linear equations.

Example: $(x_1, x_2) = (7, 8)$ is a solution to the system

$$\begin{aligned} 10x_1 + 7x_2 &= 126 \\ 5x_1 + 11x_2 &= 123 \end{aligned}$$

Example: $(-2, 5, 0)$ and $(0, 4, -1)$ are both solutions to

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

Example: For arbitrary values s_1 and s_2

$$\begin{aligned} x_1 &= s_2 - s_1 + 1 \\ x_2 &= s_2 + s_1 + 2 \\ x_3 &= s_1 \\ x_4 &= s_2 \end{aligned}$$

is a solution to

$$\begin{aligned} x_1 - 2x_2 + 3x_3 + x_4 &= -3 \\ 2x_1 - x_2 + 3x_3 - x_4 &= 0 \end{aligned}$$

We can write the set as

$$\{(s_2 - s_1 + 1, s_2 + s_1 + 2, s_1, s_2) \mid s_1, s_2 \in \mathbb{K}\} \subseteq \mathbb{K}^4$$

In some cases we may end up with a set of solutions that is empty. For example the system

$$\begin{aligned} x_1 - x_2 + 3x_3 + x_4 &= 3 \\ x_1 - x_2 + 3x_3 + x_4 &= 0 \end{aligned}$$

has no solution (the solution set is empty). In this case we say the system is *inconsistent*. Observe that in the above system of linear equation each equation on its own is consistent but the system of linear equations has no solution. The reason is that the equations have solutions

$$\begin{aligned} S_1 &= \{(3 + s_1 - 3s_2 + s_3, s_1, s_2, s_3) \mid s_1, s_2, s_3 \in \mathbb{K}\} \subseteq \mathbb{K}^4 \\ S_2 &= \{(t_1 - 3t_2 + t_3, t_1, t_2, t_3) \mid t_1, t_2, t_3 \in \mathbb{K}\} \subseteq \mathbb{K}^4 \end{aligned}$$

respectively, but $S_1 \cap S_2 = \emptyset$.

Definition 7. *[inconsistent consitent] A system of linear equations is* inconsistent *if it has no solutions; otherwise it is* consistent.

Example: here is another inconsistent system of linear equations.

$$\begin{array}{rrrrrrcr}
2x_1 & +2x_2 & +3x_3 & +3x_4 & & +x_6 & = & 3 \\
 & & x_3 & +2x_4 & +x_5 & & = & -2 \\
 & & & 7x_4 & +x_5 & & = & 1 \\
 & & & & -x_5 & +2x_6 & = & 0 \\
 & & & & & 0 & = & 1
\end{array}$$

Definition 8 (homogeneous system of equations)**.** *A system of equations is called* homogeneous *if each equation is homogeneous.*

Every system of linear equations S in variables $\mathbf{X} = \{x_1, \ldots, x_n\}$

$$\begin{array}{rcl}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\
 & \vdots & \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m
\end{array}$$

has a corresponding homogeneous system of linear equation, which is in the same set of variable and obtained by changing the constant of every equation to zero namely,

$$\begin{array}{rcl}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & 0 \\
 & \vdots & \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & 0.
\end{array}$$

Example: For the non-homogeneous system of linear equations.

$$\begin{array}{rrrrrrcr}
2x_1 & +2x_2 & +3x_3 & +3x_4 & & +x_6 & = & 3 \\
 & & x_3 & +2x_4 & +x_5 & & = & -2 \\
 & & & 7x_4 & +x_5 & & = & 1 \\
 & & & & -x_5 & +2x_6 & = & 0 \\
 & & & & & 0 & = & 1
\end{array}$$

the corresponding homogeneous system of linear equations is

$$\begin{array}{rrrrrrcr}
2x_1 & +2x_2 & +3x_3 & +3x_4 & & +x_6 & = & 0 \\
 & & x_3 & +2x_4 & +x_5 & & = & 0 \\
 & & & 7x_4 & +x_5 & & = & 0 \\
 & & & & -x_5 & +2x_6 & = & 0 \\
 & & & & & 0 & = & 0
\end{array}$$

Remark: any homogeneous system of linear equations and its corresponding homogeneous system of linear equations are the same.

1.3 Equivalent systems

Definition 9. *Two systems of equations S_1 and S_2 are* equivalent *if they have the same set of solutions.*

Example: For example

$$S_1 : \begin{cases} x_1 - 2x_2 + 3x_3 + x_4 & = & -3 \\ x_1 - 2x_2 + 3x_3 + x_4 & = & -3 \\ 2x_1 - x_2 + 3x_3 - x_4 & = & 0 \end{cases} , \quad S_2 : \begin{cases} x_1 - 2x_2 + 3x_3 + x_4 & = & -3 \\ 2x_1 - x_2 + 3x_3 - x_4 & = & 0 \\ 2x_1 - x_2 + 3x_3 - x_4 & = & 0 \end{cases}$$

and

$$S_3 : \begin{cases} 2x_1 - x_2 + 3x_3 - x_4 & = & 0 \\ x_1 - 2x_2 + 3x_3 + x_4 & = & -3 \\ 2x_1 - x_2 + 3x_3 - x_4 & = & 0 \end{cases}$$

have the same set of solutions:

$$\begin{aligned} x_1 &= s_2 - s_1 + 1 \\ x_2 &= s_2 + s_1 + 2 \\ x_3 &= s_1 \\ x_4 &= s_2 \end{aligned}$$

Here $s_1, s_2 \in \mathbb{K}$. The following system while closely related is not equivalent to the above system of linear equations as it has a unique solution:

$$\begin{aligned} x_1 - 2x_2 + 3x_3 + x_4 &= -3 \\ 2x_1 - x_2 + 3x_3 - x_4 &= 0 \\ x_3 &= 0 \\ x_4 &= 0 \end{aligned}$$

given by the singleton

$$(1, 2, 0, 0)$$

Consider the following two systems of linear equations:

$$S_x : \begin{cases} x_1 - 2x_2 + 3x_3 + x_4 & = & -3 \\ 4x_1 - 5x_2 + 9x_3 + x_4 & = & -6 \end{cases}$$

and

$$S_y : \begin{cases} y_1 - 2y_2 + 3y_3 + y_4 & = & -3 \\ 4y_1 - 5y_2 + 9y_3 + y_4 & = & -6 \end{cases}$$

These are trivially equivalent as S_x is in variables $\{x_1, x_2, x_3, x_4\}$ and S_y is in $\{y_1, y_2, y_3, y_4\}$. Thus as far as set of solutions is concerned (which is what we are interested in), is carried by the coefficients and the constants of each equation; the variables' labels are irrelevant. We will therefore represent system of linear equations via *matrices*.

1.4 Matrices and vectors

Definition 10 (matrix)**.** *An $m \times n$* matrix $A = \{a_{ij}\}_{1 \le i \le m, 1 \le j \le n}$ *is a rectangular array of numbers with m rows and n columns. Each number in the matrix is called an* entry*.*

Examples: The matrix M_1 is a 3×5, the matrix M_2 is a 5×3 matrix:

$$M_1 = \begin{pmatrix} 4 & 1 & 0 & -1 & 9 \\ 5 & 2 & 1 & 7 & 1 \\ -3 & 0 & 5 & 8 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 8 & 10 & -6 \\ 2 & 4 & 0 \\ 0 & 2 & 10 \\ -2 & 14 & 16 \\ 18 & 2 & 2 \end{pmatrix}$$

The following is *not* a matrix:

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

Definition 11 (equal matrices)**.** *Let*

$$A = \{a_{ij}\}_{1 \le i \le m, 1 \le j \le n} \qquad \textit{and} \qquad B = \{b_{ij}\}_{1 \le i \le r, 1 \le j \le p}$$

be two matrices. We say that $A = B$ if $m = r$, $n = p$ and for all $1 \le i \le m, 1 \le j \le n$ we have that $a_{ij} = b_{ij}$

Example: The matrix M_1 is equal to itself but it is not equal to the matrix given by

$$\begin{pmatrix} 4 & 1 & 0 & -1 & 9 & 0 \\ 5 & 2 & 1 & 7 & 1 & 0 \\ -3 & 0 & 5 & 8 & 1 & 0 \end{pmatrix}$$

Definition 12 (square matrix)**.** *An $m \times n$ matrix is called* square of order m *if $m = n$.*

Example: The matrix M_3 is a 3×3, i.e. a square matrix of order three:

$$M_3 = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 7 \\ 0 & 5 & 8 \end{pmatrix}$$

Definition 13 (row vector)**.** *An $1 \times n$ matrix is called a* row vector*. The entries in a vector are also called* components*.*

Example: a row vector with five components

$$\vec{r} = (4,\ 1,\ 0,\ -1,\ 9)$$

Definition 14 (column vector)**.** *An $m \times 1$ matrix is called a* column vector.[2]

Example: a column vector with five components

$$\vec{c} = \begin{pmatrix} 4 \\ 1 \\ 0 \\ -1 \\ 9 \end{pmatrix}$$

Remark: $\vec{c} \neq \vec{r}$.

1.5 Representations of system of linear equations

Definition 15 (coefficient matrix and augmented matrix of a system of linear equations)**.** *Let S be a system of linear equations in $\{x_1, \ldots, x_n\}$ given by*

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

Let A be the $m \times n$ matrix with entry ij equal a_{ij} and b is a column vector with ith component b_i. The matrix A is called the (coefficient) matrix of the system*. The* augmented matrix *of the system $(A|b)$ is the $m \times (n+1)$ matrix with ij entry equal a_{ij} if $j \leq n$ and b_i otherwise. That is*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \ddots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad (A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ & \ddots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Example: For the system of linear equations

$$\begin{array}{ccccccc} x_1 & +3x_2 & +3x_3 & +2x_4 & +x_5 & = & 7 \\ 3x_1 & +9x_2 & -6x_3 & +4x_4 & +3x_5 & = & -7 \\ 2x_1 & +6x_2 & -4x_3 & +2x_4 & +2x_5 & = & -4 \end{array}$$

[2]Often instead of a column vector we will say only a vector.

the matrix of the system is

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix}$$

and its *augmented matrix* is

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 3 & 9 & -6 & 4 & 3 & -7 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right)$$

Example: For the system of linear equations

$$\begin{array}{cccccrr} 2y_1 & +6y_2 & -4y_3 & +2y_4 & +2y_5 & = & -2 \\ & & & y_4 & & = & -1 \\ & & 5y_3 & +y_4 & & = & 9 \end{array}$$

the matrix is

$$\begin{pmatrix} 2 & 6 & -4 & 2 & 2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 5 & 1 & 0 \end{pmatrix}$$

the augmented matrix is

$$\left(\begin{array}{ccccc|c} 2 & 6 & -4 & 2 & 2 & -2 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 5 & 1 & 0 & 9 \end{array}\right)$$

Example: For the system of linear equations

$$\begin{aligned} \frac{4}{7}x_1 + 6x_2 + 9x_3 &= 0 \\ 4x_2 - 12x_3 &= 0 \\ 12x_1 &= 0 \\ 0 &= 1 \end{aligned}$$

the matrix of the system is

$$\begin{pmatrix} \frac{4}{7} & 6 & 9 \\ 0 & 4 & -12 \\ 12 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and the augmented matrix is

$$\left(\begin{array}{ccc|c} \frac{4}{7} & 6 & 9 & 0 \\ 0 & 4 & -12 & 0 \\ 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Definition 16 (vector representation of a system of linear equations)**.** *Let S be a system of linear equations in $\{x_1, \dots, x_n\}$ given by*

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

The vector representation of the system is

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Example: For the system

$$\begin{array}{rrrrrcr} x_1 & +3x_2 & +3x_3 & +2x_4 & +x_5 & = & 7 \\ 3x_1 & +9x_2 & -6x_3 & +4x_4 & +3x_5 & = & -7 \\ 2x_1 & +6x_2 & -4x_3 & +2x_4 & +2x_5 & = & -4 \end{array}$$

the corresponding vector representation is

$$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} x_1 + \begin{pmatrix} 3 \\ 9 \\ 6 \end{pmatrix} x_2 + \begin{pmatrix} 3 \\ -6 \\ -4 \end{pmatrix} x_3 + \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} x_4 + \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} x_5 = \begin{pmatrix} 7 \\ -7 \\ -4 \end{pmatrix}$$

Example: For the system of linear equations

$$\begin{array}{rcl} \frac{4}{7}x_1 + 6x_2 + 9x_3 & = & 0 \\ 4x_2 - 12x_3 & = & 0 \\ 12x_1 & = & 0 \\ 0 & = & 1 \end{array}$$

the vector form is

$$\begin{pmatrix} \frac{4}{7} \\ 0 \\ 12 \\ 0 \end{pmatrix} x_1 + \begin{pmatrix} 6 \\ 4 \\ 0 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} 9 \\ -12 \\ 0 \\ 0 \end{pmatrix} x_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

1.6 Echelon form and back substitution

Definition 17 (leading (basic) variable)**.** *In each row of a system, the first variable with a nonzero coefficient is the row's* leading (basic) variable*.*

Definition 18 (Echelon form)**.** *A system is in* Echelon form *if each leading variable is to the right of the leading variable in the row above it, except for the leading variable in the first row, and any all-zero rows are at the bottom.*

Note: The above form we well call Upper Triangular form as many software tools call Echelon form the Reduced Echelon form which will be discussed a bit later in this text.

Example: The system of linear equations

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 3 \\ -x_1 - 3x_2 + 3x_3 &= 1 \end{aligned}$$

is *not* in Echelon Form. It has overall one leading variable namely x_1.

Example: The system of linear equations

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ 3x_3 + 4x_4 &= 1 \\ 0 &= 0 \end{aligned}$$

is in Echelon form and x_1 and x_3 are the leading variables. There are no other leading variables.

Definition 19 (free variable)**.** *The non-leading variable in an Echelon form are called* free variables*.*

In Echelon form a system can be easily solved (say using a computer) using *back-substitution*, which is essentially going from the bottom equation and moving up. At each stage the current set of solutions is intersected with the set of solution that satisfy the equation that is processed. To solve a single equation with more that one variable assigns a parameter to each free variable and represent the leading variable via the assigned parameters. Consider

$$\begin{array}{rrrrrrcr} 2x_1 & +2x_2 & +3x_3 & +3x_4 & & +x_6 & = & 3 \\ & & x_3 & +2x_4 & +x_5 & & = & -2 \\ & & & 7x_4 & +x_5 & & = & 1 \\ & & & & -x_5 & +2x_6 & = & 0 \end{array}$$

From the last equation x_5 is leading and x_6 is free variable. We assign x_6 a parameter, say $s_1 \in \mathbb{K}$. Then

$$\begin{aligned} x_5 &= 2s_1 \\ x_6 &= s_1 \end{aligned}$$

after rewriting the last equation and substituting x_6 with its parameter. We now move to the equation above the last one. Here x_4 is leading variable. We have computed x_5 and therefore writing x_4 in terms of x_5 and x_6 we obtain

$$\begin{aligned} x_4 &= \frac{1}{7} - \frac{2}{7}s_1 \\ x_5 &= 2s_1 \\ x_6 &= s_1 \end{aligned}$$

Moving one equation up with leading variable x_3, we now know the values of x_4, x_5 and x_6. Expressing x_3 with the knowledge we have so far we obtain

$$\begin{aligned} x_3 &= -2 - \frac{2}{7} + \frac{4}{7}s_1 - 2s_1 = -\frac{16}{7} - \frac{10}{7}s_1 \\ x_4 &= \frac{1}{7} - \frac{2}{7}s_1 \\ x_5 &= 2s_1 \\ x_6 &= s_1 \end{aligned}$$

It remains to look at the first equation. Here we see a new free variable that we did not encounter so far, namely x_2 and x_1 is leading variable. Just as we did for x_6 we assign a parameter for x_2 say s_2. Note that the values of s_2 and s_1 are independent from each other, so now we have

$$\begin{aligned} x_2 &= s_2 \\ x_3 &= -\frac{16}{7} - \frac{10}{7}s_1 \\ x_4 &= \frac{1}{7} - \frac{2}{7}s_1 \\ x_5 &= 2s_1 \\ x_6 &= s_1; \end{aligned}$$

and with this information we can also express x_1 as

$$\begin{aligned} 2x_1 &= 3 - 2s_2 - 3\left(-\frac{16}{7} - \frac{10}{7}s_1\right) - 3\left(\frac{1}{7} - \frac{2}{7}s_1\right) - s_1 \\ &= \frac{66}{7} - 2s_2 + \frac{29}{7}s_1 \end{aligned}$$

to get the complete set of solutions as

$$\begin{aligned} x_1 &= \frac{66}{14} - s_2 + \frac{29}{14}s_1 \\ x_2 &= s_2 \\ x_3 &= -\frac{16}{7} - \frac{10}{7}s_1 \\ x_4 &= \frac{1}{7} - \frac{2}{7}s_1 \\ x_5 &= 2s_1 \\ x_6 &= s_1; \end{aligned}$$

The set of solution can be represented in the so called *vector form*

$$\left\{ \begin{pmatrix} \frac{66}{14} \\ 0 \\ \frac{16}{7} \\ \frac{1}{7} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} s_2 + \begin{pmatrix} \frac{29}{14} \\ 0 \\ -\frac{10}{7} \\ -\frac{2}{7} \\ 2 \\ 1 \end{pmatrix} s_1 \mid s_1, s_2 \in \mathbb{K} \right\}$$

Observe that even though the system given by

$$\begin{aligned} 2x_1 + x_2 + 3x_3 + 3x_4 + x_6 + 7x_7 + 9x_8 &= 1 \\ 5x_1 + 9x_2 + 5x_3 + 3x_4 + x_5 + x_6 + 7x_7 + 9x_8 &= 2 \\ 3x_1 + 7x_2 + 3x_3 + 7x_4 + x_5 + x_6 + 6x_7 + 9x_8 &= 3 \\ 3x_1 + x_3 + 3x_4 + x_6 + 7x_7 + 9x_8 &= 4 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{cccccccc|c} 2 & 1 & 3 & 3 & 0 & 1 & 7 & 9 & 1 \\ 5 & 9 & 5 & 3 & 1 & 1 & 7 & 9 & 2 \\ 3 & 7 & 3 & 7 & 1 & 1 & 6 & 9 & 3 \\ 3 & 0 & 1 & 3 & 0 & 1 & 7 & 9 & 4 \end{array} \right)$$

seems simpler in terms constants involved, it is nevertheless harder to solve compared to

$$\begin{aligned} x_1 - \frac{1}{157}x_5 + \frac{28}{157}x_6 + \frac{217}{157}x_7 + \frac{252}{157}x_8 &= \frac{192}{157} \\ x_2 + \frac{23}{157}x_5 - \frac{16}{157}x_6 - \frac{124}{157}x_7 - \frac{144}{157}x_8 &= -\frac{20}{157} \\ x_3 - \frac{12}{157}x_5 + \frac{22}{157}x_6 + \frac{341}{314}x_7 + \frac{198}{157}x_8 &= -\frac{259}{314} \\ x_4 + \frac{5}{157}x_5 + \frac{17}{157}x_6 + \frac{185}{314}x_7 + \frac{153}{157}x_8 &= \frac{121}{314} \end{aligned}$$

with augmented matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -\frac{1}{157} & \frac{28}{157} & \frac{217}{157} & \frac{252}{157} & \frac{192}{157} \\ 0 & 1 & 0 & 0 & \frac{23}{157} & -\frac{16}{157} & -\frac{124}{157} & -\frac{144}{157} & -\frac{20}{157} \\ 0 & 0 & 1 & 0 & -\frac{12}{157} & \frac{22}{157} & \frac{341}{314} & \frac{198}{157} & -\frac{259}{314} \\ 0 & 0 & 0 & 1 & \frac{5}{157} & \frac{17}{157} & \frac{185}{314} & \frac{153}{157} & \frac{121}{314} \end{pmatrix}.$$

Indeed back substitution applies to easily to the second system but no general procedure can be applied to the first. Consequently, to solve a system of linear equation we need a method to transform it into an equivalent system that is in Echelon form.

1.7 Row operations as matrix multiplication

Recall the system

$$\begin{array}{rrrrrcr} x_1 & +3x_2 & +3x_3 & +2x_4 & +x_5 & = & 7 \\ 3x_1 & +9x_2 & -6x_3 & +4x_4 & +3x_5 & = & -7 \\ 2x_1 & +6x_2 & -4x_3 & +2x_4 & +2x_5 & = & 4 \end{array}$$

From the system we can obtain new equations by combining existing equations: for example by adding 2 times Equation 1 to Equation 3 we obtain

$$\begin{array}{rrrrrcr} 4x_1 & +12x_2 & +2x_3 & +6x_4 & +4x_5 & = & 18 \end{array}$$

by adding negative 2 times Equation 1 to Equation 2 we obtain

$$\begin{array}{rrrrrcr} x_1 & +3x_2 & -12x_3 & & +x_5 & = & -21 \end{array}$$

by adding Equation 2 to itself we obtain

$$\begin{array}{rrrrrcr} 6x_1 & +18x_2 & -12x_3 & +8x_4 & +6x_5 & = & -14 \end{array}$$

by adding negative 2 times Equation 2 to Equation 3 we obtain

$$\begin{array}{rrrrrcr} -4x_1 & -12x_2 & +8x_3 & -6x_4 & -4x_5 & = & 18 \end{array}$$

by adding negative 2 times Equation 1 to Equation 3 we obtain

$$\begin{array}{rrcr} -10x_3 & -2x_4 & = & -10 \end{array}$$

As a result we obtain a new system of linear equations

$$\begin{array}{rrrrrcr} 4x_1 & +12x_2 & +2x_3 & +6x_4 & +4x_5 & = & 18 \\ x_1 & +3x_2 & -12x_3 & & +x_5 & = & -21 \\ 6x_1 & +18x_2 & -12x_3 & +8x_4 & +6x_5 & = & -14 \\ -4x_1 & -12x_2 & +8x_3 & -6x_4 & -4x_5 & = & 18 \\ & & -10x_3 & -2x_4 & & = & -10 \end{array}$$

We want to encode such transformation so that there is an easy and convenient way to work just with the (augmented) matrices of the system of linear equation. Note that adding negative 2 times Equation 1 to Equation 3, adding negative 2 times Equation 1 to Equation 3 and adding negative 2 times Equation 2 to Equation 3, appear to have the same constants $(-2, 1)$ we want to distinguish between those so rather than saying we add negative two times Equation 1 to Equation 3 we will say add negative two times Equation 1 to zero times Equation 2 to one times Equation 3. Then we can distinguish between the combinations $(-2, 0, 1)$, $(-2, 1, 0)$ and $(0, -2, 1)$. Thus we can encode the above transformation in a matrix

$$rowcomb = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -2 & 1 \\ -2 & 0 & 1 \end{pmatrix}.$$

In matrix $rowcomb$ the first raw indicates that we take 2 times Equation 1, add zero times Equation 2, add 1 times Equation 3. We can have as many such combinations as we want. To identify on which system we apply those operations we write

$$\begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 3 & 9 & -6 & 4 & 3 & -7 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right)$$

The result is a system of linear equation with matrix

$$\left(\begin{array}{rrrrr|r} 4 & 12 & 2 & 6 & 4 & 10 \\ 1 & 3 & -12 & 0 & 1 & -21 \\ 6 & 18 & -12 & 8 & 6 & -14 \\ -4 & -12 & 8 & -6 & -4 & 10 \\ 0 & 0 & -10 & -2 & 0 & -18 \end{array}\right)$$

We call this procedure matrix multiplication:

$$\begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 3 & 9 & -6 & 4 & 3 & -7 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right) = \left(\begin{array}{rrrrr|r} 4 & 12 & 2 & 6 & 4 & 10 \\ 1 & 3 & -12 & 0 & 1 & -21 \\ 6 & 18 & -12 & 8 & 6 & -14 \\ -4 & -12 & 8 & -6 & -4 & 10 \\ 0 & 0 & -10 & -2 & 0 & -18 \end{array}\right)$$

Observe that the first row of the right side is $2 \times Eqn1 + 0 \times Eqn2 + 1 \times Eqn3$ in particular it is expressed as a linear combination of the rows the augmented matrix of the original system of linear equations. The restriction that we place when multiplying matrices $AB = C$ in that case is that we require that the number of columns of A equals the number of rows of B. In other words we can obtain as many new equations as we want (number of rows of A is the same as the number of rows of C). We may have as many variables as we want (number of columns of B equals number of columns of C). The result C has its rows represented as linear combinations of the rows of B.

1.7.1 Matrix operations

Previously, we considered row operations and worked towards representing the combinations of equations via matrices. We required multiplication of an equation with a scalar (recall an equation corresponds to a row in the augmented matrix, thus it is simply a row vector). We needed addition of two equations. And lastly we multiplied matrices to obtain the result. Multiplying equation with a scalar requires that each coefficient is multiplied by the said scalar. Generalized to matrices we get

Definition 20 (scalar matrix multiplication)**.** *Let $A = \{a_{ij}\}$ be a $m \times n$ matrix and c be a constant. We define cA as the $m \times n$ with entries $\{ca_{ij}\}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.*

Example:

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

When adding two equations we added constants in front of variables, which in vector forms is simple component-wise addition (note the row vectors must have the same number of components). Generalizing to matrix addition we get

Definition 21 (matrix addition)**.** *Let* $A = \{a_{ij}\}$ *and* $B = \{b_{ij}\}$ *be a two* $m \times n$ *matrices. Define* $A + B$ *as the* $m \times n$ *matrix with entries* $\{a_{ij} + b_{ij}\}$ *for* $1 \leq i \leq m$ *and* $1 \leq j \leq n$*.*

Example:

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 3 & 9 & -6 & 4 \\ 2 & 6 & -4 & 2 \end{pmatrix} + \begin{pmatrix} 3 & 0 & 3 & 1 \\ -8 & 1 & -6 & 0 \\ 4 & 1 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 6 & 3 \\ -5 & 10 & -12 & 4 \\ 6 & 7 & -2 & 7 \end{pmatrix}$$

Lastly, from the way we wrote the transformation via the augmented matrices we can define

Definition 22 (matrix multiplication)**.** *Let*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ & \vdots & & \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{pmatrix},$$

Define $C = AB$ *by* $c_{uv} = \sum_{i=1}^{n} a_{ui} b_{iv}$*.*

Example:

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & 0 & 4 & 0 \\ 2 & 3 & -4 & 2 & 4 \end{pmatrix} \begin{pmatrix} 3 & -8 \\ 0 & 1 \\ 3 & -6 \\ 1 & 0 \\ 2 & -3 \end{pmatrix} = \begin{pmatrix} 16 & -26 \\ 13 & -15 \\ 4 & -1 \end{pmatrix}$$

Remark: we emphasize the difference between scalar multiplication and matrix multiplication: if A is a $n \times m$ matrix with $n > 1$ then αA is defined for any scalar α but $[\alpha]A$ is *not* defined, where $[\alpha]$ denotes the 1×1 matrix with entry α.

Definition 23 (linear combination of vectors)**.** *A* linear combination *of vectors $\vec{v}_1 \ldots \vec{v}_n$ is an expression of the form*

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n,$$

where a_i's are coefficients and belong to $\mathbb{K}$.

In matrix multiplication $AB = C$ we have that rows of C are linear combinations of rows of B; also columns of C are linear combinations of columns of A.

Example: consider the matrix multiplication

$$\begin{pmatrix} 0 & 5 & -2 & -1 \\ -1 & 2 & 3 & 0 \\ 2 & -4 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ -2 & 2 \\ -1 & -2 \\ 2 & -3 \end{pmatrix} = \begin{pmatrix} -10 & 17 \\ -8 & 1 \\ 16 & -30 \end{pmatrix}$$

for the second column of the result we have

$$(-3)\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} + 2\begin{pmatrix} 5 \\ 2 \\ -4 \end{pmatrix} + (-2)\begin{pmatrix} -2 \\ 3 \\ 2 \end{pmatrix} + (-3)\begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 17 \\ 1 \\ -30 \end{pmatrix}$$

for the third row of the result we have

$$2\,(1,\, -3) + (-4)\,(-2,\, 2) + 2\,(-1,\, -2) + 4\,(2,\, -3) = (16,\, -30)$$

Matrix representation of SLE: From the vector form of a system of linear equation we can write its matrix form $A\vec{x} = \vec{b}$ where A is the coefficient matrix of the system.

Example: for the system of lineary equation

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= -3 \\ 3x_1 + 9x_2 - 6x_3 + 4x_4 &= 2 \\ 2x_1 + 6x_2 - 4x_3 + 2x_4 &= 5 \end{aligned}$$

' which has vector form

$$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} x_1 + \begin{pmatrix} 3 \\ 9 \\ 6 \end{pmatrix} x_2 + \begin{pmatrix} 3 \\ -6 \\ -4 \end{pmatrix} x_3 + \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} x_4 = \begin{pmatrix} -3 \\ 2 \\ 5 \end{pmatrix}$$

has matrix representation

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 3 & 9 & -6 & 4 \\ 2 & 6 & -4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \\ 5 \end{pmatrix}$$

and very often the vector of variables is not written in a long i.e. the matrix representation is

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 3 & 9 & -6 & 4 \\ 2 & 6 & -4 & 2 \end{pmatrix} \vec{x} = \begin{pmatrix} -3 \\ 2 \\ 5 \end{pmatrix}.$$

Properties. Scalar matrix multiplication and matrix matrix addition has some similarities with the usual addition and multiplication with real numbers. In terms of matrix addition, whenever two matrices can be added the following are satisfied:

- $A + B = B + A$
- $A + (B + C) = (A + B) + C$

Furthermore there is a *zero matrix* $0_{m\times n}$, whose entries are all zeroes such that $A + 0_{m\times n} = 0_{m\times n} + A = A$. Similarly, for every matrix A there is a matrix B such that $A + B = 0_{m\times n}$. Typically B is denoted via $-A$.

Note: It is important to observe that to add two matrices they must have the same number of rows and columns. However, to multiply two matrices the number of columns of the first matrix must equal the number of rows of the second matrix. So it is possible to add two 2×5 matrices, but it is not possible to multiply them. It is possible to multiply a 2×5 by a 5×2 matrix. In that sense it is not always possible to multiply a matrix A with itself (e.g., if A is a 2×5 matrix). However it is possible to multiply A by its transpose.

Definition 24 (transpose)**.** *Let $A = \{a_{ij}\}$ be a $m \times n$ matrix. The* transpose *of A denoted by A^T is a $n \times m$ matrix $\{a^T_{uv}\}$, where $a^T_{uv} = a_{vu}$ for all $1 \le u \le n$ and $1 \le v \le m$.*

Example: the matrix

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & 0 & 4 & 0 \\ 2 & 3 & -4 & 2 & 4 \end{pmatrix}$$

has transpose

$$\begin{pmatrix} 1 & 3 & 2 \\ 3 & 9 & 3 \\ 3 & 0 & -4 \\ 2 & 4 & 2 \\ 1 & 0 & 4 \end{pmatrix}.$$

Special class of matrices: the following definitions are often encountered in practice:

Definition 25 (square matrix)**.** *An $n \times n$ matrix is called a* square *matrix. Often it is called* square matrix of order *n.*

Example: a square matrix of order three

$$\begin{pmatrix} 24 & 38 & 7 \\ 38 & 106 & 41 \\ 7 & 41 & 49 \end{pmatrix}$$

and order five

$$\begin{pmatrix} 14 & 36 & -5 & 18 & 9 \\ 36 & 99 & -3 & 48 & 15 \\ -5 & -3 & 25 & -2 & -13 \\ 18 & 48 & -2 & 24 & 10 \\ 9 & 15 & -13 & 10 & 17 \end{pmatrix}.$$

Definition 26 (lower triangular matrix)**.** *Let $A = \{a_{ij}\}$ be an $n \times n$ matrix. If for all $1 \leq i < j \leq n$ we have that $a_{ij} = 0$ then A is called a* lower triangular *matrix.*

In other words all elements above the diagonal are all zero.

Example:

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 4 & -5 & 2 & 0 \\ 3 & 0 & 1 & -1 \end{pmatrix}$$

Definition 27 (upper triangular matrix)**.** *Let $A = \{a_{ij}\}$ be an $n \times n$ matrix. If for all $1 \leq j < i \leq n$ we have that $a_{ij} = 0$ then A is called a* upper triangular *matrix.*

In other words all elements below the diagonal are all zero.

Example:

$$\begin{pmatrix} 2 & -1 & 3 & 1 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

Definition 28 (diagonal matrix)**.** *Let $A = \{a_{ij}\}$ be an $n \times n$ matrix. If A is both lower and upper triangular then A is called a* diagonal *matrix.*

Example:

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

Definition 29 (scalar matrix)**.** *Let $A = \{a_{ij}\}$ be an $n \times n$ matrix. If A is diagonal and all its diagonal entries are equal then A is called a* scalar *matrix.*

Example:

$$\begin{pmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$$

Definition 30 (identity matrix)**.** *The scalar matrix with diagonal entries equal to* 1 *is called the* identity *matrix and denoted by* I.

Example:

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Properties of matrix multiplication:

- $AB \neq BA$, in fact one of the multiplications may not even exist! Matrix multiplication is *not* commutative in general.
- $A(BC) = (AB)C$
- $A(B + C) = AB + AC$ – multiplication is left distributive over addition.
- $(B+C)D = BD+CD$ – multiplication is right distributive over addition.
- $AI = IA = A$ – left and right I may not be the same

With real numbers if $ab = 0$ then it must be the case that $a = 0$ or $b = 0$. However with matrices this is not the case. If for a non-zero square matrix A there exists a non-zero matrix B such that $AB = 0_{m\times n}$ then A is called a *zero divisor*.

Example Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ then $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Both A and B happen to be zero divisors.

Example

$$A = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 2 \\ -6 & -4 \end{pmatrix} \qquad AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Both A and B are zero divisors.

Definition 31 (inverse and invertible matrix)**.** *A square matrix A is called* invertible *if there is matrix B such that $BA = I$. The matrix B is called the* inverse *of A and denoted by A^{-1}.*

Example: the inverse of matrix

$$A = \begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix}$$

is matrix

$$A^{-1} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ -1 & 3 & 1 \end{pmatrix}.$$

Theorem 2. *If $AB = I$ then $BA = I$.*

Proof. Assume by contradiction that $BA \neq I$. Then multiplying both sides on the right by B we obtain $(BA)B \neq IB$. The right hand side due to associativity of matrix multiplication is $B(AB)$ and since $AB = I$ we have $(BA)B = B(AB) = BI = B$. On the left hand side we have $IB = B$ thus we obtain $B \neq B$ a contradiction. Therefore $BA = I$. □

Theorem 3. *If A is invertible matrix then the inverse of A is unique.*

Proof. Suppose $AC = I$ and $AD = I$. By Theorem 2 we have $AD = DA = I$. Since D is a right inverse of A, then D is also a left inverse of A. Left and right inverses are equal thus the right inverse of A which is C equals the left inverse which is D, so $C = D$ completing the argument.

$$D = DI = D(AC) = (DA)C = IC = C$$

□

Theorem 4. *If for a matrix A there exists matrices B and C such that $AB = I$ and $AC = 0$ then $C = 0$, where 0 is the zero matrix.*

Proof. If $AB = I$ then by Theorem 2 $BA = I$.

$$AC = 0_{k\times k} \Rightarrow B(AC) = B0_{k\times k} \Rightarrow (BA)C = 0_{k\times k} \Rightarrow IC = 0_{k\times k} \Rightarrow C = 0_{k\times k}$$

□

1.8 Gauss' method

We can solve a system of linear equations using back-substitution, but only if the system is in Echelon form. The next theorem gives us a way to transform a system of linear equations into Echelon form.

Theorem 5. *[Gauss' method] If a linear system S is changed to another S' by one of these operations:*

1. *an equation is swapped with another*

2. *an equation has both sides multiplied by a non-zero constant*

3. *an equation is replaced by the sum of itself and a multiple of another*

then the two system of equations have the same set of solutions.

Proof. 1. Exercise

2. Homework question

3. We have to show two the set of solution for S is the same as the set of solution for S'. To show equality of two sets we show that every element from the first set is also an element of the second set and then show that every element of the second set is an element of the first set.

   $S \subseteq S'$: Suppose $s_1, s_2, \ldots, s_n$ is a solution to S then

$$S : \left[\begin{array}{lllllll} a_{11}s_1 & +a_{12}s_2 & +\cdots & +a_{1n}s_n & = & b_1 \\ a_{21}s_1 & +a_{22}s_2 & +\cdots & +a_{2n}s_n & = & b_2 \\ \vdots & & & & & \\ a_{m1}s_1 & +a_{m2}s_2 & +\cdots & +a_{mn}s_n & = & b_m \end{array} \right.,$$

   and therefore in S' for each $k \neq j$ we have

$$a_{k1}s_1 + a_{k2}s_2 + \cdots + a_{kn}s_n = b_k.$$

   It remains to verify that

$$a'_{j1}s_1 + a'_{j2}s_2 + \cdots + a'_{jn}s_n = b'_j. \tag{1.1}$$

   Without loss of generality suppose S' was obtained from S by adding c times equation t to equation j in S. That is $a'_{ji} = a_{ji} + ca_{ti}$ for $1 \leq i \leq n$, and $b'_j = b_j + cb_t$. For Equation 1.1 we then have

$$\begin{aligned} a'_{j1}s_1 + a'_{j2}s_2 + \cdots + a'_{jn}s_n &= (a_{j1} + ca_{t1})s_1 + (a_{j2} + ca_{t2})s_2 + \cdots \\ &\qquad \cdots + (a_{jn} + ca_{tn})s_n \\ &= a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n \\ &\qquad +ca_{t1}s_1 + ca_{t2}s_2 + \cdots + ca_{tn}s_n \\ &= a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n \\ &\qquad +c(a_{t1}s_1 + a_{t2}s_2 + \cdots + a_{tn}s_n) \\ &= b_j + cb_t = b'_j. \end{aligned}$$

   Thus $s_1, \ldots, s_n$ is also a solution to S'.

$S \supseteq S'$ Conversely, suppose $s'_1, s'_2, \ldots, s'_n$ is a solution to S' then

$$S' : \left[\begin{array}{lllllll} a_{11}s'_1 & +a_{12}s'_2 & +\cdots & +a_{1n}s'_n & = & b_1 \\ \vdots & & & & & \\ a'_{j1}s'_1 & +a'_{j2}s'_2 & +\cdots & +a'_{jn}s'_n & = & b'_j \\ \vdots & & & & & \\ a_{m1}s'_1 & +a_{m2}s'_2 & +\cdots & +a_{mn}s'_n & = & b_m \end{array} \right. ,$$

and therefore in S for each $k \neq j$ we have

$$a_{k1}s_1 + a_{k2}s_2 + \cdots + a_{kn}s_n = b_k.$$

It remains to verify that

$$a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n = b_j. \tag{1.2}$$

Without loss of generality suppose S' was obtained from S by adding c times equation t to equation j in S. That is $a'_{ji} = a_{ji} + ca_{ti}$ for $1 \leq i \leq n$ or $a_{ji} = a'_{ji} - ca_{ti}$, and $b'_j = b_j + cb_k k$ implying $b_j = b'_j - cb_t$. For Equation 1.2 we then have

$$\begin{aligned} a_{j1}s'_1 + a_{j2}s'_2 + \cdots + a_{jn}s'_n &= (a'_{j1} - ca_{t1})s'_1 + (a'_{j2} - ca_{t2})s'_2 + \\ &\quad + \cdots + (a'_{jn} - ca_{tn})s'_n \\ &= a'_{j1}s'_1 + a'_{j2}s'_2 + \cdots + a'_{jn}s'_n \\ &\quad - ca_{t1}s'_1 - ca_{t2}s'_2 - \cdots - ca_{tn}s'_n \\ &= a'_{j1}s'_1 + a'_{j2}s'_2 + \cdots + a'_{jn}s'_n \\ &\quad - c(a_{t1}s'_1 + a_{t2}s'_2 + \cdots + a_{tn}s'_n) \\ &= b'_j - cb_t = b_j. \end{aligned}$$

Thus $s'_1, \ldots, s'_n$ is also a solution to S.

Therefore S and S' have the same set of solutions.

□

Definition 32 (elementary row operations)**.** *The* elementary row operations, *(also* row operations, Gaussian operations*) are*

1. *row swapping*
2. *rescaling (multiplication with a non-zero constant)*
3. *row combinations (adding a multiple of another row)*

Reducing to Echelon form: a system of linear equations in Echelon form is easy to solve, by the above mentioned back substitution. One procedure to obtain an equivalent system of linear equations to ensure by swapping equation that the leading variable of the first equation call it x is not to the right of any leading variable of the remaining equations. Then by adding suitable multiples of the first equation to the other equations so that the coefficient in front of x in the all equations except the first one are zero. The same procedure is recursively applied to the system linear equation that is obtained by removing the first equation. Until a single equation remains at which stage the procedure terminates. It is important to note that all the operations that were performed are elementary row operations. Thus by Theorem 5 all the system of equations are equivalent (that is they have the same set of solutions).

1.8.1 Gauss' method example:

recall the system

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ 3x_1 + 9x_2 - 6x_3 + 4x_4 + 3x_5 &= -7 \\ 2x_1 + 6x_2 - 4x_3 + 2x_4 + 2x_5 &= -4 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 3 & 9 & -6 & 4 & 3 & -7 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right)$$

Step 1**:** add negative three times equation one to equation two to get

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ -15x_3 - 2x_4 &= -28 \\ 2x_1 + 6x_2 - 4x_3 + 2x_4 + 2x_5 &= -4 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 3 & 9 & -6 & 4 & 3 & -7 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right).$$

Step 2**:** add negative two times equation one to equation three to get

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ -15x_3 - 2x_4 &= -28 \\ -10x_3 - 2x_4 &= -18 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & -10 & -2 & 0 & -18 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right)$$

Step 3: add $\dfrac{-2}{3}$ times equation two to equation three to get

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ -15x_3 - 2x_4 &= -28 \\ -\frac{2}{3}x_4 &= \frac{2}{3} \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & -\frac{2}{3} & 0 & \frac{2}{3} \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & -10 & -2 & 0 & -18 \end{array}\right).$$

The system is in Echelon form and can be solved using back substitution. The solution set is

$$\begin{aligned} x_1 &= 3 - 3t_1 - t_2 \\ x_2 &= t_1 \\ x_3 &= 2 \\ x_4 &= -1 \\ x_5 &= t_2 \end{aligned}$$

in vector form

$$\left\{ \begin{pmatrix} 3 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} t_1 + \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} t_2 \mid t_1, t_2 \in \mathbb{K} \right\}$$

The above Echelon form is suitable for back substitution (i.e. it has a “inverted stair” shape), but we can continue with elementary row operations to obtain the *Reduced Echelon form* which allows for even easier way to identify solution (in many ways finding solution with back substitution involves in a convoluted way getting the Reduced Echelon form).

Step 4: scale equation three by $\dfrac{-3}{2}$ to get

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ -15x_3 - 2x_4 &= -28 \\ x_4 &= -1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{3}{2} \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & -\frac{2}{3} & 0 & \frac{2}{3} \end{array}\right).$$

Step 5: add two times equation three to equation two to get

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ -15x_3 &= -30 \\ x_4 &= -1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & 0 & 0 & -30 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right)$$

Step 6: scale equation two by $\dfrac{-1}{15}$ to get

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ x_3 &= 2 \\ x_4 &= -1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{15} & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & 0 & 0 & -30 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right).$$

Step 7: add negative three times equation two to equation one to get

$$\begin{aligned} x_1 + 3x_2 + 2x_4 + x_5 &= 1 \\ x_3 &= 2 \\ x_4 &= -1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 0 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right).$$

Step 8**:** add negative two times equation three to equation one to get

$$\begin{aligned} x_1 + 3x_2 + x_5 &= 3 \\ x_3 &= 2 \\ x_4 &= -1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{ccccc|c} 1 & 3 & 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccccc|c} 1 & 3 & 0 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right).$$

Reduced Echelon form The above equation is in Reduced Echelon form, meaning that

1. the system is in Echelon form;
2. the coefficients in front of leading variables is each one;
3. every leading variable appears in exactly one equation.

With the Reduced Echelon form performing Back Substitution requires no extra computation. Many software tool that provide routines for Gaussian elimination when called upon produce in fact the Reduced Echelon Form. Some linear algebra texts identify the Reduced Echelon form with the Echelon form. From now on when we say Echelon form we will mean Reduced Echelon form. In rare cases when the inverted stair shape is asked for or needed it will be made explicit.

As said with the Reduced Echelon form it is easier to identify the solution. It is also useful when you need to solve multiple SLEs with the same coefficient matrix.

1.8.2 Inconsistent system of linear equation

Consider the following system of linear equations

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 3 \\ -x_1 - 3x_2 + 3x_3 &= 2 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 5 & 3 \\ -1 & -3 & 3 & 0 & 2 \end{array}\right)$$

Step 1**:** add -2 times equation one to equation two

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ 3x_3 + x_4 &= 1 \\ -x_1 - 3x_2 + 3x_3 &= 2 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ -1 & -3 & 3 & 0 & 2 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 5 & 3 \\ -1 & -3 & 3 & 0 & 2 \end{array}\right).$$

Step 2**:** add 1 times equation one to equation three

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ 3x_3 + x_4 &= 1 \\ 6x_3 + 2x_4 &= 3 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 6 & 2 & 3 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right) \left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ -1 & -3 & 3 & 0 & 2 \end{array}\right).$$

Step 3**:** add -2 times equation two to equation three

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ 3x_3 + x_4 &= 1 \\ 0 &= 1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{array}\right) \left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 6 & 2 & 3 \end{array}\right).$$

Step 4**:** scale equation two by $\dfrac{1}{3}$

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ x_3 + \frac{1}{3}x_4 &= \frac{1}{3} \\ 0 &= 1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

Step 5**:** add -3 times equation two to equation one

$$\begin{aligned} x_1 + 3x_2 + x_4 &= 0 \\ x_3 + \frac{1}{3}x_4 &= \frac{1}{3} \\ 0 &= 1 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

The system is in Echelon form. Since the last equation has no solution, the system of linear equations has no solution, in other words it is inconsistent, equivalently the solution set is *empty* or the solutions set equals the empty set denoted by $\emptyset$.

The fact that system of linear equations is inconsistent becomes apparent after Step 3, however, we as mentioned earlier we will be reducing all our system of linear equations to Reduced Echelon form.

1.8.3 Another consistent example

Consider the set of equations in Section 1.8.2 but with different constants

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 3 \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 5 \\ -x_1 - 3x_2 + 3x_3 &= -5 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 3 \\ 2 & 6 & 9 & 5 & 5 \\ -1 & -3 & 3 & 0 & -5 \end{array}\right).$$

Step 1**:** add -2 times equation one to equation two

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 3 \\ 3x_3 + x_4 &= -1 \\ -x_1 - 3x_2 + 3x_3 &= -5 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 3 & 1 & -1 \\ -1 & -3 & 3 & 0 & -5 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 3 \\ 2 & 6 & 9 & 5 & 5 \\ -1 & -3 & 3 & 0 & -5 \end{array}\right).$$

Step 2**:** add 1 times equation one to equation three

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 3 \\ 3x_3 + x_4 &= -1 \\ 6x_3 + 2x_4 &= -2 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 3 & 1 & -1 \\ 0 & 0 & 6 & 2 & -2 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 3 & 1 & -1 \\ -1 & -3 & 3 & 0 & -5 \end{array}\right).$$

Step 3**:** add -2 times equation two to equation three

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 3 \\ 3x_3 + x_4 &= -1 \\ 0 &= 0 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 3 & 1 & -1 \\ 0 & 0 & 6 & 2 & -2 \end{array}\right).$$

Step 4**:** scale equation two by $\frac{1}{3}$

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 3 \\ x_3 + \frac{1}{3}x_4 &= -\frac{1}{3} \\ 0 &= 0 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Step 5**:** add -3 times equation two to equation one

$$\begin{aligned} x_1 + 3x_2 + x_4 &= 4 \\ x_3 + \frac{1}{3}x_4 &= -\frac{1}{3} \\ 0 &= 0 \end{aligned}$$

represented as matrix multiplication

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 4 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 3 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The system is in Echelon form and can be solved using back substitution. The solution set is

$$\begin{aligned} x_1 &= 4 - 3t_1 - t_2 \\ x_2 &= t_1 \\ x_3 &= -\frac{1}{3} - \frac{1}{3}t_2 \\ x_4 &= t_2 \end{aligned}$$

in vector form

$$\left\{ \begin{pmatrix} 4 \\ 0 \\ -\frac{1}{3} \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} t_1 + \begin{pmatrix} -1 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix} t_2 \mid t_1, t_2 \in \mathbb{K} \right\}.$$

1.8.4 System of linear equations with shared matrix

The system of linear equations from Section 1.8.2

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 3 \\ -x_1 - 3x_2 + 3x_3 &= 2 \end{aligned}$$

and the system of linear equations from Section 1.8.3

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 3 \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 5 \\ -x_1 - 3x_2 + 3x_3 &= -5 \end{aligned}$$

have the same (coefficient) matrix

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix}.$$

Reduction to Echelon form does not depend on the constants of the equations so we can combine the two reductions from the previous sections into single reduction. It can be efficiently done using augmented matrices, namely we combine the augmented matrix of the first system of linear equations

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 5 & 3 \\ -1 & -3 & 3 & 0 & 2 \end{array}\right)$$

with the augmented matrix of the second system of linear equations

$$\left(\begin{array}{rrrr|r} 1 & 3 & 3 & 2 & 3 \\ 2 & 6 & 9 & 5 & 5 \\ -1 & -3 & 3 & 0 & -5 \end{array}\right)$$

to obtain an augmented matrix for two system of linear equations sharing the same matrix namely,

$$\left(\begin{array}{rrrr|rr} 1 & 3 & 3 & 2 & 1 & 3 \\ 2 & 6 & 9 & 5 & 3 & 5 \\ -1 & -3 & 3 & 0 & 2 & -5 \end{array}\right).$$

To the left of the vertical separator we have the matrix of the system of linear equation, to its left every column corresponds to a different system of linear equations whose constants are given by that column.

Step 1**:** add -2 times equation one to equation two represented as matrix multiplication

$$\left(\begin{array}{rrrr|rr} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ -1 & -3 & 3 & 0 & 2 & -5 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{rrrr|rr} 1 & 3 & 3 & 2 & 1 & 3 \\ 2 & 6 & 9 & 5 & 3 & 5 \\ -1 & -3 & 3 & 0 & 2 & -5 \end{array}\right).$$

Step 2: add 1 times equation one to equation three represented as matrix multiplication

$$\left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ 0 & 0 & 6 & 2 & 3 & -2 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ -1 & -3 & 3 & 0 & 2 & -5 \end{array}\right).$$

Step 3: add -2 times equation two to equation three represented as matrix multiplication

$$\left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ 0 & 0 & 6 & 2 & 3 & -2 \end{array}\right).$$

Step 4: scale equation two by $\dfrac{1}{3}$ represented as matrix multiplication

$$\left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

Step 5: add -3 times equation two to equation one represented as matrix multiplication

$$\left(\begin{array}{cccc|cc} 1 & 3 & 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right) = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

From here we can extract the augmented matrices of the two system of linear equation. For the first one we have

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

which corresponds to

$$\begin{aligned} x_1 + 3x_2 + x_4 &= 0 \\ x_3 + \frac{1}{3}x_4 &= \frac{1}{3} \\ 0 &= 1 \end{aligned}$$

For the second one we have augmented matrix

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 4 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

which corresponds to

$$\begin{aligned} x_1 + 3x_2 + x_4 &= 4 \\ x_3 + \frac{1}{3}x_4 &= -\frac{1}{3} \\ 0 &= 0 \end{aligned}$$

1.8.5 System of linear equations with unique solution

Consider the system of linear equations

$$\begin{aligned} -x_2 + x_3 &= 1 \\ 2x_1 + 3x_2 - 2x_3 &= 0 \\ x_1 + 2x_2 - x_3 &= 0 \end{aligned}$$

the system of linear equations

$$\begin{aligned} -x_2 + x_3 &= 0 \\ 2x_1 + 3x_2 - 2x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 0 \end{aligned}$$

and the system of linear equations

$$\begin{aligned} -x_2 + x_3 &= 0 \\ 2x_1 + 3x_2 - 2x_3 &= 0 \\ x_1 + 2x_2 - x_3 &= 1 \end{aligned}$$

They share the same matrix and therefore we will solve them simultaneously. Constructing the combined augmented matrix we get

$$\left(\begin{array}{rrr|rrr} 0 & -1 & 1 & 1 & 0 & 0 \\ 2 & 3 & -2 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right).$$

Step 1**:** swap equation one and equation three, given as matrix multiplication

$$\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 2 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \left(\begin{array}{rrr|rrr} 0 & -1 & 1 & 1 & 0 & 0 \\ 2 & 3 & -2 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right).$$

Step 2**:** add -2 times equation one to equation two, given as matrix multiplication

$$\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & -2 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 2 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right).$$

Step 3**:** scale equation two by -1, given as matrix multiplication

$$\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & -2 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right).$$

Step 4**:** add 1 times equation two to equation three, given as matrix multiplication

$$\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 & 2 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right)\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right).$$

Step 5**:** add 1 times equation three to equation one, given as matrix multiplication

$$\left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & -1 & 3 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 & 2 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 & 2 \end{array}\right).$$

Step 6**:** add -2 times equation two to equation one, given as matrix multiplication

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 & 2 \end{array}\right) = \left(\begin{array}{rrr} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & -1 & 3 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 & 2 \end{array}\right).$$

From here we can extract the augmented matrices of the two system of linear equation. For the first one we have

$$\left(\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

which corresponds to

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 0 \\ x_3 &= 1 \end{aligned}$$

The system is in Echelon form and can be solved using back substitution. In this particular case the above description coincides with the solution set (which is a singleton). In vector form

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

For the second one we have

$$\left(\begin{array}{ccc|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{array} \right)$$

which corresponds to

$$\begin{aligned} x_1 &= 1 \\ x_2 &= -1 \\ x_3 &= -1 \end{aligned}$$

The system is in Echelon form and can be solved using back substitution. In this particular case the above description coincides with the solution set (which is a singleton). In vector form

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right\}.$$

For the third one we have

$$\left(\begin{array}{ccc|r} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

which corresponds to

$$\begin{aligned} x_1 &= -1 \\ x_2 &= 2 \\ x_3 &= 2 \end{aligned}$$

The system is in Echelon form and can be solved using back substitution. In this particular case the above description coincides with the solution set (which is a singleton). In vector form

$$\left\{ \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} \right\}$$

1.8.6 Observations

Parallelization Consider Step 2 from Section 1.8.4

$$\left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ 0 & 0 & 6 & 2 & 3 & -2 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|cc} 1 & 3 & 3 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 & 1 & -1 \\ -1 & -3 & 3 & 0 & 2 & -5 \end{array}\right)$$

We have that

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 6 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ -1 & -3 & 3 & 0 \end{array}\right)$$

and

$$\begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & -5 \end{pmatrix}$$

In general if a matrix $M = (C \mid D)$ is multiplied with a matrix A that is if you compute AM the computation can be parallelized by computing AM as $(AC \mid AD)$.

Associativity Matrix multiplication is associative operation. Consider Step 1 and Step 2 from Section 1.8.5

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 0 & 1 \\ 2 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \left(\begin{array}{ccc|ccc} 0 & -1 & 1 & 1 & 0 & 0 \\ 2 & 3 & -2 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right)$$

and

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & -2 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 0 & 1 \\ 2 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 \end{array}\right)$$

Using the previous observation for the right side of the separator (the constants of the three system of linear equations) we have

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

For the left side of the separator we have

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

In general if we perform Gaussian operations on multiple system of linear equations as in the previous section at any stage if we have a combined augmented matrix $(U \mid V)$ and an initial shared matrix B we have the relations

$$U = VB$$

by applying the above procedure to the first n-steps (i.e. apply associativity of matrix multiplication to first multiply the Gaussian operations into a single matrix before applying it to the matrix of the system of linear equations).

Here is another example: consider Step 5 from Section 1.8.5 we have a resulting combined augmented matrix

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & -1 & 3 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 & 2 \end{array}\right)$$

given the initial shared matrix

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

the following relation holds:

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

The above observation is a basis for computing inverses of matrices.

1.8.7 Reduced Echelon form and inverse

To compute the inverse of matrix

$$\begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix}$$

establish the following three system of linear equation: the first one

$$\begin{aligned} 5x_1 - 8x_2 + x_3 &= 1 \\ 3x_1 - 5x_2 + x_3 &= 0 \\ -4x_1 + 7x_2 - x_3 &= 0 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{ccc|c} 5 & -8 & 1 & 1 \\ 3 & -5 & 1 & 0 \\ -4 & 7 & -1 & 0 \end{array}\right)$$

the second one

$$\begin{aligned} 5x_1 - 8x_2 + x_3 &= 0 \\ 3x_1 - 5x_2 + x_3 &= 1 \\ -4x_1 + 7x_2 - x_3 &= 0 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{rrr|r} 5 & -8 & 1 & 0 \\ 3 & -5 & 1 & 1 \\ -4 & 7 & -1 & 0 \end{array}\right)$$

and the third one

$$\begin{aligned} 5x_1 - 8x_2 + x_3 &= 0 \\ 3x_1 - 5x_2 + x_3 &= 0 \\ -4x_1 + 7x_2 - x_3 &= 1 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{rrr|r} 5 & -8 & 1 & 0 \\ 3 & -5 & 1 & 0 \\ -4 & 7 & -1 & 1 \end{array}\right).$$

The combined augmented matrix is

$$\left(\begin{array}{rrr|rrr} 5 & -8 & 1 & 1 & 0 & 0 \\ 3 & -5 & 1 & 0 & 1 & 0 \\ -4 & 7 & -1 & 0 & 0 & 1 \end{array}\right)$$

Performing Gaussian eliminations

Step 1**:**

$$\left(\begin{array}{rrr|rrr} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{3}{5} & 1 & 0 \\ -4 & 7 & -1 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ -\frac{3}{5} & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \left(\begin{array}{rrr|rrr} 5 & -8 & 1 & 1 & 0 & 0 \\ 3 & -5 & 1 & 0 & 1 & 0 \\ -4 & 7 & -1 & 0 & 0 & 1 \end{array}\right)$$

Step 2**:**

$$\left(\begin{array}{rrr|rrr} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{3}{5} & 1 & 0 \\ -4 & 7 & -1 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{rrr} 1 & 0 & 0 \\ -\frac{3}{5} & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \left(\begin{array}{rrr|rrr} 5 & -8 & 1 & 1 & 0 & 0 \\ 3 & -5 & 1 & 0 & 1 & 0 \\ -4 & 7 & -1 & 0 & 0 & 1 \end{array}\right)$$

Step 3:

$$\left(\begin{array}{ccc|ccc} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{3}{5} & 1 & 0 \\ 0 & \frac{3}{5} & -\frac{1}{5} & \frac{4}{5} & 0 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{3}{5} & 1 & 0 \\ -4 & 7 & -1 & 0 & 0 & 1 \end{array}\right)$$

Step 4:

$$\left(\begin{array}{ccc|ccc} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{3}{5} & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{3}{5} & 1 & 0 \\ 0 & \frac{3}{5} & -\frac{1}{5} & \frac{4}{5} & 0 & 1 \end{array}\right)$$

Step 5:

$$\left(\begin{array}{ccc|ccc} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & 0 & -\frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{3}{5} & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right)$$

Step 6:

$$\left(\begin{array}{ccc|ccc} 5 & -8 & 0 & 2 & -3 & -1 \\ 0 & -\frac{1}{5} & 0 & -\frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 5 & -8 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{5} & 0 & -\frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right)$$

Step 7:

$$\left(\begin{array}{ccc|ccc} 5 & -8 & 0 & 2 & -3 & -1 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 5 & -8 & 0 & 2 & -3 & -1 \\ 0 & -\frac{1}{5} & 0 & -\frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right)$$

Step 8:

$$\left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 10 & 5 & 15 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} 1 & 8 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 5 & -8 & 0 & 2 & -3 & -1 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right)$$

Step 9:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 1 & 3 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} \frac{1}{5} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 10 & 5 & 15 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right)$$

Ultimately, using the second observation from the previous section we have

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ -1 & 3 & 1 \end{pmatrix} \begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix}$$

thus the desired inverse is

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

To compute the inverse of a matrix then augment the matrix with the identity matrix, reduce the resulting system of linearly equations to Reduced Echelon form, the set of solutions (column-wise) is then the inverse of the matrix.

In the above procedure every step is a Gaussian operation. From the previous section the observations imply that the inverse is the product of Gaussian operations. For the example above it means that

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ -1 & 3 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 8 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{5} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This is an algorithm to compute the inverse of a matrix if one exists. It also proves the following theorem (subject to a few technical details).

Theorem 6. *If A is an invertible matrix then A can be written as a product of elementary matrices.*

1.9 Homogeneous and particular solutions

Example: Recall the system with augmented matrix

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 &= 7 \\ 3x_1 + 9x_2 - 6x_3 + 4x_4 + 3x_5 &= -7 \\ 2x_1 + 6x_2 - 4x_3 + 2x_4 + 2x_5 &= -4 \end{aligned}$$

with augmented matrix

$$\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 3 & 9 & -6 & 4 & 3 & -7 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right)$$

after Gaussian elimination the Reduced Echelon form is

$$\left(\begin{array}{ccccc|c} 1 & 3 & 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right)$$

whose solution set can be described in vector form as

$$\left\{ \begin{pmatrix} 3 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} s_1 + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -3 \end{pmatrix} s_2 \mid s_1, s_2 \in \mathbb{K} \right\}$$

In matrix form the system of linear equations is

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 7 \\ -7 \\ -4 \end{pmatrix}$$

Using matrix operations that were defined earlier, substitute the vectors from the vector form of the solution in the matrix equation $A\vec{x} = \vec{b}$.

For the vector without any parameters

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ -7 \\ -4 \end{pmatrix}$$

For the vector in front of s_1

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

For the vector in front of s_2

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

In the vector representation of the solution if a vector $\vec{h}$ satisfies $A\vec{h} = \vec{0}$, where in this case $\vec{0}$ is the vector with all components equal zero then it is part of the homogeneous solution and if a vector $\vec{p}$ satisfies $A\vec{p} = \vec{b}$ then it is a particular solution. In general the set of solution to $A\vec{x} = \vec{b}$ is given by a particular solution plus the set of homogeneous solutions.

Here is another homogeneous solution

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \left[(2) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} + (-3) \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -3 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 0 \\ 0 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Here is another particular solution

$$\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \left[\begin{pmatrix} 3 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix} + (2) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} + (-3) \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -3 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ -3 \\ 2 \\ -1 \\ 7 \end{pmatrix} = \begin{pmatrix} 7 \\ -7 \\ -4 \end{pmatrix}$$

For the set of solutions

$$\left\{ \underbrace{\begin{pmatrix} 3 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix}}_{\text{particular solution}} + \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} s_1 + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -3 \end{pmatrix} s_2}_{\text{homogeneous solution}} \mid s_1, s_2 \in \mathbb{K} \right\}$$

Theorem 7. *Any system of linear equations has a description of the solution set in the form*

$$\{\vec{p} + c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \mid c_1, \ldots, c_k \in \mathbb{K}\} \tag{1.3}$$

where $\vec{p}$ is any particular solution and where the number of vectors $\vec{\beta}_1, \ldots, \vec{\beta}_k$ equals the number of free variables that the system has after a Gaussian reduction to Echelon form.

Argument. The proof of the above theorem relies on the following lemmas.

Lemma 1. *For any homogeneous linear system there exist vectors $\vec{\beta}_1, \ldots, \vec{\beta}_k$ such that the solution set of the system is*

$$\{c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \mid c_1, \ldots, c_k \in \mathbb{K}\} \tag{1.4}$$

where k is the number of free variables in an Echelon form version of the system.

Proof. Apply Gauss's Method to get to Echelon form. Observe that the system of linear equations has at least one solution (the tuple of all zeroes) and it does not contain equations of the form $0 = b$ for a non-zero constant b.

If the system of linear equation has only $0 = 0$ type equations then all variables a free variables and any tuple is a solution thus the lemma holds.

If the system of linear equations has both equations with non-zero coefficients and some $0 = 0$ equations, ignore the $0 = 0$ equations since their solution contains all tuples.

By induction we will verify that each leading variable can be expressed in terms of free variables. That implies the lemma since the free variables can be used as parameters and the $\vec{\beta}$'s are the vectors of coefficients of those free variables.

For the base step consider the bottom-most equation

$$a_{m,\ell_m}x_{\ell_m} + a_{m,\ell_m+1}x_{\ell_m+1} + \cdots + a_{m,n}x_n = 0 \tag{1.5}$$

where $a_{m,\ell_m} \neq 0$. (x_{ℓ_m} is the leading variable in row m.) At the bottom row any variables after the leading one are free. For this equation the result hold by Theorem 1 by setting the non-zero coefficient index to ℓ_m.

Assume by induction the statement holds for the bottom-most t rows, with $0 \leq t < m-1$, the leading variable can be expressed in terms of the free ones. It remains to verify that it then also holds for the $(m-(t+1))$-th equation.

Take each leading variable in a lower equation $x_{\ell_m}, \ldots, x_{\ell_{m-t}}$ and substitute its expression in terms of free variables.

Since the system is in Echelon form all such leading variables have larger index than the leading variable the $(m-(t+1))$-th equation. As a result it has a leading term of

$$a_{m-(t+1),\ell_{m-(t+1)}}x_{\ell_{m-(t+1)}}$$

with

$$a_{m-(t+1),\ell_{m-(t+1)}} \neq 0,$$

and the rest of the left hand side is a linear combination of free variables. Rearranging by moving the free variables to the right side and dividing by $a_{m-(t+1),\ell_{m-(t+1)}}$ expresses this equation's leading variable $x_{\ell_{m-(t+1)}}$ in terms of the free variables.

Thus by induction the result follows. $\square$

Lemma 2. *Let $\vec{p} = (p_1, \ldots, p_n)$ be any particular solution to a system of linear equations, then the solution set of the system of linear equations is the set:*

$$\mathcal{S} = \{\vec{p} + \vec{h} \mid \vec{h} \text{ satisfies the corresponing homogeneous system}\}$$

Proof. Let $\vec{s} = (s_1, \ldots, s_n)$ be a solution to the system of linear equations.

Consider $\vec{\vec{h}} = \vec{p} - \vec{s}$, substitute in the i'th equation in the corresponding homogeneous system of linear equations to obtain

$$\begin{aligned} & a_{i,1}(s_1 - p_1) + \cdots + a_{i,n}(s_n - p_n) \\ = \quad & (a_{i,1}s_1 + \cdots + a_{i,n}s_n) - (a_{i,1}p_1 + \cdots + a_{i,n}p_n) \\ = \quad & d_i - d_i = 0 \end{aligned}$$

Thus $\vec{s} = \vec{p} - \vec{\vec{h}}$ and therefore any solution is in the set $\mathcal{S}$

Conversely, take $\vec{p} + \vec{h}$, where $\vec{h}$ solves the associated homogeneous system. For an equation i in the system of linear equations the following holds:

$$\begin{aligned} & a_{i,1}(p_1 + h_1) + \cdots + a_{i,n}(p_n + h_n) \\ = \quad & (a_{i,1}p_1 + \cdots + a_{i,n}p_n) + (a_{i,1}h_1 + \cdots + a_{i,n}h_n) \\ = \quad & d_i + 0 = d_i \end{aligned}$$

so any vector $\vec{p} + \vec{h}$ is a solution to the linear system of equations. $\square$

A homogeneous system of linear equations always has at least one solution. If there are free variables then such a homogeneous system of linear equations has infinitely many solutions. Thus if a system of linear equations has a solution it either has a unique solutions or infinitely many solution; it may have no solutions at all. The following table summarizes the possibilities.

		number of solutions of the homogeneous system	
		one	*infinitely many*
particular solution exists?	*yes*	unique solution	infinitely many solutions
	no	no solutions	no solutions

1.9.1 Zero equals zero and number of solutions

The system of linear equations

$$\begin{array}{rrcl} x_1 & +3x_2 & = & 5 \\ 4x_1 & +12x_2 & = & k \end{array}$$

has Reduced Echelon form

$$\begin{array}{rrcr} x_1 & +3x_2 & = & 5 \\ 0x_1 & +0x_2 & = & k - 20 \end{array}$$

If $k = 20$ we have infinitely many solutions (what are they). If $k \neq 20$ we have no solutions. So does $0 = 0$ tell us that we have infinitely many solutions always? Consider the following system of linear equation

$$\begin{array}{rrrrcr} x_1 & +3x_2 & +0x_3 & +0x_4 & = & 5 \\ 4x_1 & +12x_2 & +0x_3 & +0x_4 & = & 20 \\ 0x_1 & +0x_2 & +x_3 & +3x_4 & = & 5 \\ 0x_1 & +0x_2 & +4x_3 & +12x_4 & = & 21 \end{array}$$

which has Reduced Echelon form

$$\begin{array}{rrrrcr} x_1 & +3x_2 & +0x_3 & +0x_4 & = & 5 \\ 0x_1 & +0x_2 & +x_3 & +3x_4 & = & 5 \\ 0x_1 & +0x_2 & +0x_3 & +0x_4 & = & 1 \\ 0x_1 & +0x_2 & +0x_3 & +0x_4 & = & 0 \end{array}$$

even though after Gaussian eliminations there is an equation $0 = 0$ the system of linear equations does not have infinitely many solutions. In fact it is inconsistent.

The system of linear equations

$$\begin{array}{rrrcr} x_1 & +x_2 & +x_3 & = & 0 \\ & x_2 & +x_3 & = & 0 \end{array}$$

has infinitely many solutions, but no Gaussian operations result in $0 = 0$. Thus the equation $0 = 0$ is not necessary for infinitely many solutions.

Remark: In general homogeneous equation have at least one solution. But other than that just by looking at the system we cannot say if it is consistent or not. In particular having more variable than equations does not guarantee infinitely many solutions. In fact it may not even be consistent.

Chapter 2

Vector spaces

2.1 Definitions and examples

Definition 33 (vector space). *A* vector space *over $\mathbb{K}$ is a non-empty set $\mathbf{V}$ consisting of* vectors *along with two operations:* vector addition *denoted by $+$ and scalar vector multiplication, such that the sum*[1] *of two vectors is also in $\mathbf{V}$ and for any scalar $c \in \mathbb{K}$ and any vector $\vec{v} \in \mathbf{V}$ we have $c\vec{v} \in \mathbf{V}$. Furthermore, the addition and scalar multiplication satisfy the following properties:*

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
2. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
3. *there is a unique* zero vector $\vec{0} \in \mathbf{V}$ *such that* $\forall \vec{v} \in \mathbf{V} : \vec{0} + \vec{v} = \vec{v}$.
4. *for each vector $\vec{v} \in \mathbf{V}$ there exist a unique vector $\vec{-v}$ such that $\vec{v} + (\vec{-v}) = \vec{0}$.*
5. *for each vector $\vec{v} \in \mathbf{V}$ we have $1\vec{v} = \vec{v}$*
6. *for each vector $\vec{v} \in \mathbf{V}$ and for all scalars α and β we have that $\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v}$*
7. $\alpha(\vec{u} + \vec{v}) = \alpha\vec{v} + \alpha\vec{u}$
8. $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$

Examples:

1. The set of complex numbers $\mathbb{C}$ over themselves $\mathbb{C}$ with standard addition and multiplication of complex numbers.
2. The set of real numbers $\mathbb{R}$ over themselves $\mathbb{R}$ with standard addition and multiplication of real numbers.

[1] In general a $\mathbf{V} \times \mathbf{V} \to \mathbf{V}$ binary operation

3. The set of matrices $\mathcal{M}_{n\times m}(\mathbb{K})$ of dimension $n \times m$ with entries in $\mathbb{K}$ with standard matrix addition and scalar matrix multiplication.

4. The set of column vectors $\mathbb{K}^n$ with n components from $\mathbb{K}$ with standard vector addition and scalar vector multiplication.

5. The plane $pl : 2X + 3Y - Z = 0$ in three dimensions with usual addition and scalar multiplication, formally

$$pl = \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mid 2X + 3Y - Z = 0 \right\}$$

   with standard vector operations.

6. Any plane in three dimension that passes through the origin i.e. any plane $pl : AX + BY + CZ = 0$ in three dimensions with usual addition and scalar multiplication, formally

$$pl = \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mid AX + BY + CZ = 0 \right\}$$

   with standard vector operations, where A, B and C can be any real values not simultaneously zero.

7. The set of all functions with domain the interval $[a, b]$ and codomain $\mathbb{R}$ denoted by $[a, b]^{\mathbb{R}}$ with standard functions addition and constant function multiplication.

8. The set of all continuous functions defined on a interval $[a, b]$ denoted by $\mathcal{C}[a, b]$ with standard functions addition and constant function multiplication.

9. The set of all sequences $\{a_n\}$ with standard operations from calculus.

10. The set of all sequences $\{a_n\}$ that have only finitely many non-zero terms with standard operations from calculus.

11. The set of all sequences $\{a_n\}$ that converge to zero that is $\{a_n\} \to 0$ with standard operations from calculus.

12. The set of all polynomials $\mathbf{P}$ with standard (calculus) operations on polynomials.

13. Polynomials of degree at most n denoted by $\mathbf{P}_n$ with standard (calculus) operations on polynomials.

14. All functions in the set $\{a \cos x + b \sin x \mid a, b \in \mathbb{R}\}$ with standard (calculus) operations on functions.

15. A set with single element z with operations $\alpha z = z$ and $z + z = z$.

Counterexamples:

1. With standard vector operations the vectors with integer components

$$\mathbb{Z}^2 = \left\{ \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \mid x_0, x_1 \in \mathbb{Z} \right\}.$$

   This set with these operations is *not* a vector space. Scalar multiplication for example

$$\begin{pmatrix} 4 \\ 1 \end{pmatrix} \in \mathbb{Z}^2,$$

   but

$$\pi \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 4\pi \\ 1\pi \end{pmatrix} \notin \mathbb{Z}^2.$$

   so scalar vector multiplication is not a function

$$\odot : \mathbb{R} \times \mathbb{Z}^2 \to \mathbb{Z}^2,$$

   its codomain is $\mathbb{R}^2$ and *not* $\mathbb{Z}^2$ that is

$$\odot : \mathbb{R} \times \mathbb{Z}^2 \to \mathbb{R}^2,$$

2. polynomials that evaluate to 1 at 3 - vector addition is not closed: adding any two polynomials that evaluate to 1 at 3 results in a polynomial that evaluates to 2 at 3

2.2 A special example

Let $\mathbf{CVS} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \in \mathbb{R} \right\}$ with the operations

$$\oplus \quad : \quad \mathbf{CVS} \times \mathbf{CVS} \to \mathbf{CVS}$$

$$\odot \quad : \quad \mathbb{R} \times \mathbf{CVS} \to \mathbf{CVS}$$

defined as

$$\vec{u} \oplus \vec{v} \quad = \quad \begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} x + s - 2 \\ y + t \end{bmatrix}$$

$$\alpha \odot \vec{u} \quad = \quad \alpha \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha x - 2\alpha + 2 \\ \alpha y \end{bmatrix}$$

Vector addition examples:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \oplus \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 100 \\ 200 \end{bmatrix} \oplus \begin{bmatrix} 1 \\ -4 \end{bmatrix} = \begin{bmatrix} 99 \\ 196 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

Scalar multiplication examples:

$$-1 \odot \begin{bmatrix} 100 \\ 200 \end{bmatrix} = \begin{bmatrix} -96 \\ -200 \end{bmatrix}$$

$$-1 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$2 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

$$0 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Verification: we verify all conditions for vector spaces:

1. the set is non-empty, for example $\begin{bmatrix} 1 \\ 4 \end{bmatrix} \in$ **CVS**

2. closure of vector addition: if $x, s \in \mathbb{R}$ then $x + s - 2 \in \mathbb{R}$; if $y, t \in \mathbb{R}$ then $y + t \in \mathbb{R}$. Therefore from the definition of $\oplus$ vector addition is closed.

3. closure of scalar multiplication: if $\alpha, x \in \mathbb{R}$ then $\alpha x - 2\alpha + 2 \in \mathbb{R}$; if $\alpha, y \in \mathbb{R}$ then $\alpha y \in \mathbb{R}$. Therefore from the definition of $\odot$ scalar multiplication is closed.

4. $\vec{u} \oplus \vec{v} = \vec{v} \oplus \vec{u}$

$$\begin{aligned} \vec{u} \oplus \vec{v} &= \begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} s \\ t \end{bmatrix} \\ &= \begin{bmatrix} x + s - 2 \\ y + t \end{bmatrix} \\ &= \begin{bmatrix} s + x - 2 \\ t + y \end{bmatrix} \\ &= \begin{bmatrix} s \\ t \end{bmatrix} \oplus \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \vec{v} \oplus \vec{u} \end{aligned}$$

5. $(\vec{u} \oplus \vec{v}) \oplus \vec{w} = \vec{u} \oplus (\vec{v} \oplus \vec{w})$

$$\begin{aligned}
(\vec{u} \oplus \vec{v}) \oplus \vec{w} &= \left(\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} s \\ t \end{bmatrix} \right) \oplus \begin{bmatrix} p \\ q \end{bmatrix} \\
&= \begin{bmatrix} x+s-2 \\ y+t \end{bmatrix} \oplus \begin{bmatrix} p \\ q \end{bmatrix} \\
&= \begin{bmatrix} (x+s-2)+p-2 \\ (t+y)+q \end{bmatrix} \\
&= \begin{bmatrix} x+(s+p-2)-2 \\ y+(t+q) \end{bmatrix} \\
&= \begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} s+p-2 \\ y+q \end{bmatrix} \\
&= \begin{bmatrix} x \\ y \end{bmatrix} \oplus \left(\begin{bmatrix} s \\ t \end{bmatrix} \oplus \begin{bmatrix} p \\ q \end{bmatrix} \right) \\
&= \vec{u} \oplus (\vec{v} \oplus \vec{w})
\end{aligned}$$

6. there is a unique *zero vector* $\vec{0} \in \mathbf{CVS}$ such that $\forall \vec{u} \in \mathbf{CVS} : \vec{0} \oplus \vec{u} = \vec{u}$. We want a vector $\vec{0} = \begin{bmatrix} A \\ B \end{bmatrix}$ such that $\vec{0} \oplus \vec{u} = \vec{u}$ for all choices of $\vec{u}$, then

$$\vec{0} \oplus \vec{u} = \begin{bmatrix} A \\ B \end{bmatrix} \oplus \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A+x-2 \\ B+y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} = \vec{u}$$

The above implies $A+x-2 = x$ meaning $A = 2$ and $B+y = B$ meaning $B = 0$. Indeed by setting $\vec{0} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ we obtain

$$\begin{aligned}
\vec{0} \oplus \vec{u} &= \begin{bmatrix} 2 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} 2+x-2 \\ 0+y \end{bmatrix} \\
&= \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \vec{u}
\end{aligned}$$

7. for each vector $\vec{u} \in \mathbf{CVS}$ there exist a unique vector $\vec{-u}$ such that $\vec{u} + (\vec{-u}) = \vec{0}$. Let $\vec{u} = \begin{bmatrix} x \\ y \end{bmatrix}$ be any vector; solve for $\vec{-u} = \begin{bmatrix} a \\ b \end{bmatrix}$ in

$$\begin{aligned}
\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \vec{0} &= \vec{u} \oplus \vec{-u} \\
&= \begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} \\
&= \begin{bmatrix} x+a-2 \\ y+b \end{bmatrix}
\end{aligned}$$

Thus $x + a - 2 = 2$ meaning that $a = -x + 4$ and $y + b = 0$ meaning that $y = -b$. Thus the unique vector satisfying the above condition is $-\vec{u} = \begin{bmatrix} -x+4 \\ -y \end{bmatrix}$

8. $1 \odot \vec{u} = \vec{u}$

$$\begin{aligned} 1 \odot \vec{u} &= 1 \odot \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} 1 \times x - 2 \times 1 + 2 \\ 1 \times y \end{bmatrix} \\ &= \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \vec{u} \end{aligned}$$

9. $\alpha \odot (\beta \odot \vec{u}) = (\alpha\beta) \odot \vec{u}$

$$\begin{aligned} \alpha \odot (\beta \odot \vec{u}) &= \alpha \odot \begin{bmatrix} \beta x - 2\beta + 2 \\ \beta y \end{bmatrix} \\ &= \begin{bmatrix} \alpha(\beta x - 2\beta + 2) - 2\alpha + 2 \\ \alpha\beta y \end{bmatrix} \\ &= \begin{bmatrix} \alpha\beta x - 2\alpha\beta + 2\alpha - 2\alpha + 2 \\ \alpha\beta y \end{bmatrix} \\ &= \begin{bmatrix} \alpha\beta x - 2\alpha\beta + 2 \\ \alpha\beta y \end{bmatrix} \\ &= (\alpha\beta) \odot \begin{bmatrix} x \\ y \end{bmatrix} \\ &= (\alpha\beta) \odot \vec{u} \end{aligned}$$

10. $\alpha \odot (\vec{u} \oplus \vec{v}) = \alpha \odot \vec{u} \oplus \alpha \odot \vec{v}$

$$\begin{aligned}
\alpha \odot (\vec{u} \oplus \vec{v}) &= \alpha \odot \left(\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} s \\ t \end{bmatrix} \right) \\
&= \alpha \odot \begin{bmatrix} x+s-2 \\ y+t \end{bmatrix} \\
&= \begin{bmatrix} \alpha(x+s-2) - 2\alpha + 2 \\ \alpha(y+t) \end{bmatrix} \\
&= \begin{bmatrix} \alpha x + \alpha s - 2\alpha - 2\alpha + 2 + 2 - 2 \\ \alpha y + \alpha t \end{bmatrix} \\
&= \begin{bmatrix} (\alpha x - 2\alpha + 2) + (\alpha s - 2\alpha + 2) - 2 \\ \alpha y + \alpha t \end{bmatrix} \\
&= \begin{bmatrix} \alpha x - 2\alpha + 2 \\ \alpha y \end{bmatrix} \oplus \begin{bmatrix} \alpha s - 2\alpha + 2 \\ \alpha t \end{bmatrix} \\
&= \alpha \odot \begin{bmatrix} x \\ y \end{bmatrix} \oplus \alpha \odot \begin{bmatrix} s \\ t \end{bmatrix} \\
&= \alpha \odot \vec{v} \oplus \alpha \odot \vec{u}
\end{aligned}$$

11. $(\alpha + \beta) \odot \vec{u} = \alpha \odot \vec{u} \oplus \beta \odot \vec{u}$

$$\begin{aligned}
(\alpha + \beta) \odot \vec{u} &= (\alpha + \beta) \odot \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} (\alpha+\beta)x - 2(\alpha+\beta) + 2 \\ (\alpha+\beta)y \end{bmatrix} \\
&= \begin{bmatrix} (\alpha x - 2\alpha + 2) + (\beta x - 2\beta + 2) - 2 \\ \alpha y + \beta y \end{bmatrix} \\
&= \begin{bmatrix} \alpha x - 2\alpha + 2 \\ \alpha y \end{bmatrix} \oplus \begin{bmatrix} \beta x - 2\beta + 2 \\ \beta y \end{bmatrix} \\
&= \alpha \odot \begin{bmatrix} x \\ y \end{bmatrix} \oplus \beta \odot \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \alpha \odot \vec{u} \oplus \beta \odot \vec{u}
\end{aligned}$$

All conditions for vector space hold therefore we have a vector space!

Remark: In this vector space $\vec{0} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, we computed that $\vec{0} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. In terms of $\vec{u} + \vec{0} = \vec{u}$ we have for example

$$\begin{bmatrix} 3 \\ -4 \end{bmatrix} \oplus \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \end{bmatrix} \neq \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

instead

$$\begin{bmatrix} 3 \\ -4 \end{bmatrix} \oplus \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

as a side observation

$$\begin{bmatrix} 3 \\ -4 \end{bmatrix} \oplus \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

the above is possible in the sense that if $\vec{z}$ denotes the vector with components that are zero, then for any vector $\vec{u}$ we have that

$$\vec{u} + (\vec{-u} + \vec{z}) = \vec{z}.$$

2.3 General results for vector spaces

Theorem 8. $\forall \vec{v} \in \mathbf{V}, 0\vec{v} = \vec{0}$.

Proof. Apply condition 8 above with $c_1 = c_2 = 0$ to get

$$0\vec{v} = (0+0)\vec{v} \quad = \quad 0\vec{v} + 0\vec{v}$$

Since $0 + 0 = 0$ the right hand side is $0\vec{v}$. Add to both sides the vector equation to $-0\vec{v}$ which exist by 4 to get

$$0\vec{v} + (-0\vec{v}) \quad = \quad 0\vec{v} + 0\vec{v} + (-0\vec{v})$$

The right hand side become $\vec{0}$

$$\vec{0} \quad = \quad 0\vec{v} + \vec{0}$$

By condition 3

$$\vec{0} \quad = \quad 0\vec{v}$$

□

Remark: In $0\vec{u} = \vec{0}$, the representation of the zero vector depends on the vector space. For the vector space **CVS** described in §2.2

$$0 \odot \begin{bmatrix} 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

likewise

$$0 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The reason is zero vector representation in that vector space.

Theorem 9. $\forall \vec{v} \in \mathbf{V}, (-1)\vec{v} = \vec{-v}$.

Proof. Observe that $0 = 1 - 1$ and apply it to the last equation

$$\vec{0} = 0\vec{v} = 1\vec{v} + (-1\vec{v})$$

By condition 5 we get

$$\vec{0} = \vec{v} + (-1\vec{v})$$

Add to both sides $\vec{-v}$

$$\vec{-v} + \vec{0} = \vec{-v} + \vec{v} + (-1\vec{v})$$

which by condition 3 for the righthandside and condition 4 for the lefthandside implies

$$\vec{-v} = \vec{0} + (-1\vec{v})$$

Again by condition 3 we have

$$\vec{-v} = -1\vec{v}$$

□

Remark: In $(-1)\vec{u} = \vec{-u}$, the representation of the additive inverse vector depends on the vector space and its operations. For the vector space **CVS** described in §2.2

$$-1 \odot \begin{bmatrix} 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \neq \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

likewise

$$-1 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The fundamental of the above reason is the special definition of vector addition and scalar multiplication.

Theorem 10. $\forall \alpha \in \mathbb{K}, \alpha\vec{0} = \vec{0}$

Proof 1. Let $\vec{u} \in V$ and $\alpha \in \mathbb{K}$. Then $\alpha\vec{u} = \alpha\vec{u}$. On the right hand side we have $\alpha\vec{u} = \alpha\vec{u} + \vec{0}$. On the left hand side using $\vec{u} + \vec{0} = \vec{u}$ we have $\alpha\vec{u} = \alpha(\vec{u} + \vec{0}) = \alpha\vec{u} + \alpha\vec{0}$. Thus

$$\alpha\vec{u} + \alpha\vec{0} = \alpha\vec{u} + \vec{0};$$

adding $-\alpha\vec{u}$ to both sides of the equation

$$\alpha\vec{u} + \alpha\vec{0} - \alpha\vec{u} = \alpha\vec{u} + \vec{0} - \alpha\vec{u} \qquad \Rightarrow \qquad \alpha\vec{0} + \vec{0} = \vec{0} + \vec{0}$$

and the desired result follows. □

Proof 2. We will use $0\vec{u} = \vec{0}$ for any vector $\vec{u}$:

$$\alpha\vec{0} = \alpha(0\vec{0}) = (\alpha 0)\vec{0} = 0\vec{0} = \vec{0}$$

□

Remark: as in the previous results $\alpha\vec{0} = \vec{0}$ refers to the zero vector in that particular vector space. For the vector space **CVS** described in §2.2

$$5 \odot \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 10 \\ 0 \end{bmatrix}$$

where as

$$5 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -8 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

2.4 Linear combinations

In the following we assume all sets of vectors are coming from the same vector space.

Definition 34. *A vector $\vec{w}$ is said to be linear combination of $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_k}$ if there exists constants $\alpha_1, \ldots, \alpha_k$ such that*

$$\vec{w} = \sum_{i=1}^{k} \alpha_i \vec{u_i} = \alpha_1 \vec{u_1} + \cdots + \alpha_k \vec{u_k}$$

The set of all linear combinations of vectors $\vec{u}_1, \ldots, \vec{u}_k$ is called *span* of these vectors. Span is often denoted by $\langle \vec{u}_1, \ldots, \vec{u}_k \rangle$. Further observations on span are available in §2.8.

Example: For the vector space $\mathbb{R}^4$ with standard operations

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

implying that $\vec{u} = \begin{pmatrix} 1 \\ 2 \\ -3 \\ -1 \end{pmatrix}$ is a linear combination of vectors

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Every vector $\vec{u}$ in $\mathbb{C}^n$ is a linear combination of $\vec{e}_1, \ldots, \vec{e}_n$ since the system of linear equations with augmented matrix

$$[I_n \mid \vec{u}]$$

always has a solutions (here I_n is the identity matrix of order n).

Remark: as in the previous section linear combinations depend on the choice of vector space and its operations. For example in the standard vector space $\mathbb{C}^2$ the equation

$$\begin{pmatrix} -2 \\ -2 \end{pmatrix} = x_1 \begin{pmatrix} 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

with corresponding augmented matrix

$$\left(\begin{array}{cc|c} 0 & 0 & -2 \\ 0 & 1 & -2 \end{array} \right)$$

has no solution and therefore the vector $\vec{w} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$ is not a linear combination of the vectors

$$\vec{u}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

However, for the vector space **CVS** described in §2.2 the equation

$$\begin{aligned} \begin{bmatrix} -2 \\ -2 \end{bmatrix} &= x_1 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus x_2 \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -2\,x_1 + 2 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} -2\,x_2 + 2 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} -2\,x_1 - 2\,x_2 + 2 \\ x_2 \end{bmatrix} \end{aligned}$$

with corresponding augmented matrix obtained by equating the vector components

$$\left(\begin{array}{cc|c} -2 & -2 & -4 \\ 0 & 1 & -2 \end{array} \right)$$

has a solution $\begin{pmatrix} 4 \\ -2 \end{pmatrix}$ which shows that

$$\begin{bmatrix} -2 \\ -2 \end{bmatrix} = 4 \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus (-2) \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

that is the vector $\begin{bmatrix} -2 \\ -2 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Another example: for the vector space with standard operations $\mathbb{C}^2$ we have

$$\begin{pmatrix} 6 \\ -5 \end{pmatrix} = 3 \begin{pmatrix} 2 \\ 0 \end{pmatrix} - 5 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

so $\begin{pmatrix} 6 \\ -5 \end{pmatrix}$ is a linear combination of $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ but for the vector space **CVS** described in §2.2 the equation

$$\begin{aligned} \begin{bmatrix} 6 \\ -5 \end{bmatrix} &= x_1 \odot \begin{bmatrix} 2 \\ 0 \end{bmatrix} \oplus x_2 \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} -2\,x_2 + 2 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} -2\,x_2 + 2 \\ x_2 \end{bmatrix} \end{aligned}$$

with corresponding augmented matrix obtained by equating the components of the vectors

$$\left(\begin{array}{cc|c} 0 & -2 & 4 \\ 0 & 1 & -5 \end{array} \right)$$

has no solution and therefore the vector $\begin{bmatrix} 6 \\ -5 \end{bmatrix}$ is not a linear combination of vectors $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Theorem 11. *If $\vec{w}$ is linear combination of a subset of $\vec{u}_1, \vec{v}_u, \ldots, \vec{u}_k$ then it is linear combination of all the vectors.*

Proof. Using Theorem 8 multiply each extra $\vec{u}_i$ with the constant zero. $\square$

Example consider the set of vectors

$$\vec{u}_1 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ -3 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 3 \end{pmatrix}, \vec{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 5 \\ 1 \\ 3 \end{pmatrix}, \vec{u}_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 6 \\ 2 \end{pmatrix}$$

we have that

$$\begin{pmatrix} 3 \\ -2 \\ 0 \\ 18 \\ 0 \end{pmatrix} = -2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 3 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 6 \\ 2 \end{pmatrix}$$

the above linear combination can be extended to all vectors via

$$\begin{pmatrix} 3 \\ -2 \\ 0 \\ 18 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ -3 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 3 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 0 \\ 5 \\ 1 \\ 3 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 6 \\ 2 \end{pmatrix}$$

Theorem 12. *If $\vec{w}$ is linear combination of $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$ and each $\vec{u}_i$ is a linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_t$ then $\vec{w}$ is a linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_t$.*

Proof. Let

$$\begin{aligned} \vec{u}_1 &= u_{11}\vec{v}_1 + u_{12}\vec{v}_2 + \cdots + u_{1t}\vec{v}_t \\ \vec{u}_2 &= u_{21}\vec{v}_1 + u_{22}\vec{v}_2 + \cdots + u_{2t}\vec{v}_t \\ &\vdots \\ \vec{u}_k &= u_{k1}\vec{v}_k + u_{k2}\vec{v}_2 + \cdots + u_{kt}\vec{v}_t \end{aligned}$$

then

$$\begin{aligned} \vec{w} &= w_1\vec{u}_1 + w_2\vec{u}_2 + \cdots + w_k\vec{u}_k \\ &= w_1 \underbrace{(u_{11}\vec{v}_1 + u_{12}\vec{v}_2 + \cdots + u_{1t}\vec{v}_t)}_{\vec{u}_1} \\ &\quad + w_2 \underbrace{(u_{21}\vec{v}_1 + u_{22}\vec{v}_2 + \cdots + u_{2t}\vec{v}_t)}_{\vec{u}_2} \\ &\quad + \cdots + w_k \underbrace{(u_{k1}\vec{v}_1 + u_{k2}\vec{v}_2 + \cdots + u_{kt}\vec{v}_t)}_{\vec{u}_k} \\ &= \underbrace{(w_1u_{11} + w_2u_{21} + \cdots + w_ku_{k1})}_{\tilde{w_1}}\vec{v}_1 \\ &\quad + \underbrace{(w_1u_{12} + w_2u_{22} + \cdots + w_ku_{k2})}_{\tilde{w_2}}\vec{v}_2 \\ &\quad + \cdots + \underbrace{(w_1u_{1t} + w_2u_{2t} + \cdots + w_ku_{kt})}_{\tilde{w_k}}\vec{v}_t \\ &= \tilde{w_1}\vec{v}_1 + \tilde{w_2}\vec{v}_2 + \cdots + \tilde{w_k}\vec{v}_k \end{aligned}$$

which establishes the result. □

Example: consider vectors

$$\vec{u}_1 = \begin{pmatrix} 1 \\ -2 \\ -12 \\ -4 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} 2 \\ -1 \\ 21 \\ 4 \end{pmatrix}, \vec{u}_3 = \begin{pmatrix} 3 \\ -3 \\ 25 \\ 4 \end{pmatrix}, \vec{u}_4 = \begin{pmatrix} -5 \\ 0 \\ -18 \\ -2 \end{pmatrix}, \vec{u}_5 = \begin{pmatrix} 4 \\ -2 \\ 18 \\ 2 \end{pmatrix}$$

For $\vec{w}$ defined as the linear combination of $\vec{u}_1, \dots, \vec{u}_5$ let

$$\vec{w} = \begin{pmatrix} -6 \\ -3 \\ -81 \\ -18 \end{pmatrix} = 4\begin{pmatrix} 1 \\ -2 \\ -12 \\ -4 \end{pmatrix} + 1\begin{pmatrix} 2 \\ -1 \\ 21 \\ 4 \end{pmatrix} + 0\begin{pmatrix} 3 \\ -3 \\ 25 \\ 4 \end{pmatrix} + 0\begin{pmatrix} -5 \\ 0 \\ -18 \\ -2 \end{pmatrix} - 3\begin{pmatrix} 4 \\ -2 \\ 18 \\ 2 \end{pmatrix}.$$

For the set of vectors

$$\vec{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 4 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}$$

the following relations hold

$$\begin{aligned}
\vec{u}_1 &= \begin{pmatrix} 1 \\ -2 \\ -12 \\ -4 \end{pmatrix} = -4 \begin{pmatrix} 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} \\
\vec{u}_2 &= \begin{pmatrix} 2 \\ -1 \\ 21 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} \\
\vec{u}_3 &= \begin{pmatrix} 3 \\ -3 \\ 25 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} \\
\vec{u}_4 &= \begin{pmatrix} -5 \\ 0 \\ -18 \\ -2 \end{pmatrix} = -2 \begin{pmatrix} 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} - 5 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} \\
\vec{u}_5 &= \begin{pmatrix} 4 \\ -2 \\ 18 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}.
\end{aligned}$$

We then have

$$\begin{aligned}
\vec{w} &= (4)\vec{u}_1 + (1)\vec{u}_2 + (0)\vec{u}_3 + (0)\vec{u}_4 + (-3)\vec{u}_5 \\
&= (4)\left((-4)\vec{v}_1 + (2)\vec{v}_2 + (1)\vec{v}_3\right) \\
&\quad +(1)\left((4)\vec{v}_1 + (1)\vec{v}_2 + (2)\vec{v}_3\right) \\
&\quad +(0)\left((4)\vec{v}_1 + (3)\vec{v}_2 + (3)\vec{v}_3\right) \\
&\quad +(0)\left((-2)\vec{v}_1 + (0)\vec{v}_2 + (-5)\vec{v}_3\right) \\
&\quad +(-3)\left((2)\vec{v}_1 + (2)\vec{v}_2 + (4)\vec{v}_3\right) \\
&= (-18)\vec{v}_1 + (3)\vec{v}_2 + (-6)\vec{v}_3
\end{aligned}$$

$$\begin{pmatrix} -6 \\ -3 \\ -81 \\ -18 \end{pmatrix} = -18 \begin{pmatrix} 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} - 6 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}$$

Theorem 13. *Let A and B be matrices such that $C = AB$ exists, then each column of C is a linear combination of the columns of A.*

Proof. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & & \ddots & \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{pmatrix}$$

If

$$C = AB = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ & & \ddots & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}$$

then for column s of matrix C, namely $\begin{pmatrix} c_{1s} \\ c_{2s} \\ \vdots \\ c_{ms} \end{pmatrix}$ we have that

$$\begin{aligned} c_{1s} &= a_{11}b_{1s} + a_{12}b_{2s} + \cdots + a_{1k}b_{ks} \\ c_{2s} &= a_{21}b_{1s} + a_{22}b_{2s} + \cdots + a_{2k}b_{ks} \\ &\vdots \\ c_{ms} &= a_{m1}b_{1s} + a_{m2}b_{2s} + \cdots + a_{mk}b_{ks} \end{aligned}$$

Writing the above equation all at once:

$$\begin{pmatrix} c_{1s} \\ c_{2s} \\ \vdots \\ c_{ms} \end{pmatrix} = b_{1s} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + b_{2s} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + b_{ks} \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix}$$

Thus the sth column of matrix C is a linear combination of the columns of A and the coefficients are the entries in row s of matrix B. $\square$

Example: given the equation

$$\begin{pmatrix} 4 & 3 & 4 \\ 12 & 12 & 18 \\ 0 & -3 & -6 \\ -8 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 3 & 3 \\ -3 & 0 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 4 & 3 & 4 \end{pmatrix}$$

for the third column of the result we have

$$\begin{pmatrix} 4 \\ 18 \\ -6 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 3 \\ -3 \\ 4 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 3 \\ 0 \\ -2 \end{pmatrix}$$

Theorem 14. *Let A and B be matrices such that $C = AB$ exists, then each row of C is a linear combination of the rows of B.*

Proof. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & & \ddots & \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{pmatrix}$$

If

$$C = AB = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ & & \ddots & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}$$

then for row r of matrix C, namely $\begin{pmatrix} c_{r1} & c_{r2} & \cdots & c_{rn} \end{pmatrix}$ we have that

$$\begin{aligned} c_{r1} &= a_{r1}b_{11} + a_{r2}b_{21} + \cdots + a_{rk}b_{k1} \\ c_{r2} &= a_{r1}b_{12} + a_{r2}b_{22} + \cdots + a_{rk}b_{k2} \\ &\vdots \\ c_{rn} &= a_{r1}b_{1n} + a_{r2}b_{2n} + \cdots + a_{rk}b_{kn} \end{aligned}$$

Writing the above equation all at once:

$$\begin{aligned} \begin{pmatrix} c_{r1} & c_{r2} & \cdots & c_{rn} \end{pmatrix} &= \left(\underbrace{a_{r1}b_{11} + a_{r2}b_{21} + \cdots + a_{rk}b_{k1}}_{c_{r1}} \quad \underbrace{a_{r1}b_{12} + a_{r2}b_{22} + \cdots + a_{rk}b_{k2}}_{c_{r2}} \quad \cdots \quad \underbrace{a_{r1}b_{1n} + a_{r2}b_{2n} + \cdots + a_{rk}b_{kn}}_{c_{rn}} \right) \\ &= a_{r1}\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \end{pmatrix} \\ &\quad + a_{r2}\begin{pmatrix} b_{21} & b_{22} & \cdots & b_{2n} \end{pmatrix} \\ &\quad + \cdots + a_{rk}\begin{pmatrix} b_{k1} & b_{k2} & \cdots & b_{kn} \end{pmatrix} \end{aligned}$$

Thus the rth row of matrix C is a linear combination of the rows of B and the coefficients are the entries in row r of matrix A. $\square$

Example: given the equation

$$\begin{pmatrix} -3 & 3 & 11 \\ -14 & -2 & 14 \\ 16 & 4 & -12 \\ -8 & -5 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 4 & -2 \\ -4 & 4 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} -3 & 0 & 4 \\ 1 & 1 & 1 \end{pmatrix}$$

for the second row of the result we have

$$(-14,\ -2,\ 14) = 4\,(-3,\ 0,\ 4) - 2\,(1,\ 1,\ 1)$$

2.5 Linear dependence and independence

In §2.4 the underlying question was given a system of linear equation, does it have a solution. The answer depends on the actual values and if affirmative it implies the system has at least one particular solution. With every system of linear equation a related question is does it have a unique solution or does it have infinitely many solutions. The answer to infinitely many versus unique solutions is given by the number of solution to the corresponding homogeneous system of linear equation. This section discusses that question in relation to vector spaces.

Definition 35. *[linear (in)dependence] Let $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ be a set of vectors. If*

$$a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_k\vec{v_k} = \vec{0} \quad \Rightarrow \quad a_1 = a_2 = \cdots = a_k = 0$$

then the vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ are called linearly independent *otherwise the are* linearly dependent*.*

Example: In $\mathbb{R}^2$ the vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are linearly independent, indeed the system of linear equations with augmented matrix

$$\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right)$$

has a unique solution $x_1 = 0$ and $x_2 = 0$.

Example: In $\mathbb{R}^3$ the vectors $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$ are linearly dependent since the system of linear equation

$$\left(\begin{array}{rrr|r} 1 & 0 & 2 & 0 \\ 0 & -2 & 2 & 0 \\ 1 & 3 & -1 & 0 \end{array}\right)$$

has a non-trivial solution $x_1 = -2$, $x_2 = 1$ and $x_3 = 1$.

Example: Consider the vector space **CVS** described in §2.2 the vectors

$$\vec{u}_1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \text{ and } \vec{u}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

are linearly independent, since the vector equation $\vec{0} = x_1\vec{u}_1 + x_1\vec{u}_2$ gives

$$\begin{aligned}
\begin{bmatrix} 2 \\ 0 \end{bmatrix} &= x_1 \odot \begin{bmatrix} 4 \\ 2 \end{bmatrix} \oplus x_2 \odot \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 2\,x_1 + 2 \\ 2\,x_1 + x_2 \end{bmatrix}
\end{aligned}$$

and the system of linear equations

$$\left(\begin{array}{cc|c} 2 & 0 & 0 \\ 2 & 1 & 0 \end{array}\right)$$

has a unique solution $x_1 = 0$ and $x_2 = 0$. However vectors

$$\vec{u}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \text{ and } \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

are linearly dependent since from

$$\begin{aligned}
\begin{bmatrix} 2 \\ 0 \end{bmatrix} &= x_1 \odot \begin{bmatrix} 2 \\ 0 \end{bmatrix} \oplus x_2 \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} -2\,x_2 + 2 \\ x_2 \end{bmatrix}
\end{aligned}$$

one obtains the system of linear equations

$$\left(\begin{array}{cc|c} 0 & -2 & 0 \\ 0 & 1 & 0 \end{array}\right)$$

which has a non-trivial solution (one such solution is $x_1 = 5$ and $x_2 = 0$).

Example: In $\mathcal{C}[a,b]$ – the vector space of continuous on a interval $[a,b]$ functions vectors are continuous functions. Let

$$\begin{aligned}
\vec{f_1} &= f_1(x) = \cos 2x \\
\vec{f_2} &= f_2(x) = \sin^2 x \\
\vec{f_3} &= f_3(x) = \cos^2 x \\
\vec{f_4} &= f_4(x) = e^x \\
\vec{f_5} &= f_5(x) = -3
\end{aligned}$$

The zero vector $\vec{0}$ is the function $z(x) = 0$, that is the constant function zero.

1. $\vec{f_1} = f_1(x)$ and $\vec{f_2} = f_2(x)$ are linearly independent, indeed consider the equation
$$\vec{0} = \alpha\vec{f_1} + \beta\vec{f_2}$$

in function form

$$z(x) = \alpha f_1(x) + \beta f_2(x)$$

In the function form the equation must hold for any $x \in \mathbb{R}$. That is, you first find α and β, and for those α and β the equation must hold for any $x \in \mathbb{R}$. If that is the case consider what happens at $x = 0$

$$z\left(0\right) = \alpha f_1\left(0\right) + \beta f_2\left(0\right) \quad \Rightarrow \quad 0 = \alpha\cos(2 \times 0) + \beta\text{sin}^2 0$$

Since $\cos(0) = 1$ and $\sin(0) = 0$ the equation implies

$$0 = \alpha + \beta 0.$$

Apply the same for $x = \frac{\pi}{4}$

$$z\left(\frac{\pi}{4}\right) = \alpha f_2\left(\frac{\pi}{4}\right) + \beta f_3\left(\frac{\pi}{4}\right) \quad \Rightarrow \quad 0 = \alpha\cos\left(2 \times \frac{\pi}{4}\right) + \beta\text{sin}^2\left(\frac{\pi}{4}\right)$$

Since $\cos\left(\frac{\pi}{2}\right) = 0$ and $\sin^2\left(\frac{\pi}{4}\right) = \dfrac{1}{2}$ the equation implies

$$0 = \alpha 0 + \frac{\beta}{2}.$$

The set of equations

$$\begin{aligned} 0 &= \alpha 0 + \frac{\beta}{2} \\ 0 &= \alpha + \beta 0 \end{aligned}$$

has unique solutions $\alpha = 0$ and $\beta = 0$.

2. The vectors $\vec{f_2} = f_2(x)$, $\vec{f_3} = f_3(x)$ and $\vec{f_5} = f_5(x)$ are linearly dependent. Consider

$$\alpha\vec{f_2} + \beta\vec{f_3} + \gamma\vec{f_5} = \vec{0}$$

since the vector equation

$$6\vec{f_2} + 6\vec{f_3} + 2\vec{f_5} = \vec{0}$$

equivalently in function form

$$6\underbrace{\sin^2(x)}_{\vec{f_2}} + 6\underbrace{\cos^2(x)}_{\vec{f_3}} + 2(\underbrace{-3}_{\vec{f_5}}) = z(x)$$

is satisfied for *any* value $x \in \mathbb{R}$.

3. The vectors $\vec{f_1} = f_1(x)$, $\vec{f_3} = f_3(x)$ and $\vec{f_4} = f_4(x)$ are linearly independent, the equation

$$\vec{0} = \alpha\vec{f_1} + \beta\vec{f_3} + \gamma\vec{f_4}$$

in function form

$$z(x) = \alpha f_1(x) + \beta f_3(x) + \gamma f_4(x)$$

evaluated at $x = 0$ implies

$$\begin{aligned} z(0) &= \alpha f_1(0) + \beta f_3(0) + \gamma f_4(0) \\ \Rightarrow 0 &= \alpha \cos(2 \times 0) + \beta \cos^2(0) + \gamma e^0 \\ \Rightarrow 0 &= \alpha + \beta + \gamma \end{aligned}$$

The same equation evaluated at $x = \pi$ implies

$$\begin{aligned} z(\pi) &= \alpha f_1(\pi) + \beta f_3(\pi) + \gamma f_4(\pi) \\ \Rightarrow 0 &= \alpha \cos(2\pi) + \beta \cos^2(\pi) + \gamma e^{\pi} \\ \Rightarrow 0 &= \alpha + \beta + \gamma e^{\pi} \end{aligned}$$

The same equation evaluated at $= \dfrac{\pi}{2}$ implies

$$\begin{aligned} z\left(\frac{\pi}{2}\right) &= \alpha f_1\left(\frac{\pi}{2}\right) + \beta f_3\left(\frac{\pi}{2}\right) + \gamma f_4\left(\frac{\pi}{2}\right) \\ \Rightarrow 0 &= \alpha \cos\left(2\frac{\pi}{2}\right) + \beta \cos^2\left(\frac{\pi}{2}\right) + \gamma e^{\frac{\pi}{2}} \\ \Rightarrow 0 &= -\alpha + \gamma e^{\frac{\pi}{2}} \end{aligned}$$

The set of equations

$$\begin{aligned} 0 &= \alpha + \beta + \gamma \\ 0 &= \alpha + \beta + \gamma e^{\pi} \\ 0 &= -\alpha + \gamma e^{\frac{\pi}{2}} \end{aligned}$$

has only one solution, namely $\alpha = 0$, $\beta = 0$ and $\gamma = 0$.

Theorem 15. *The standard basis vectors are linearly independent, in other words the columns and rows of I are linearly independent.*

Proof. Let $\vec{e_i}$ be the vector whose ith coordinate is one and the rest zero. Consider the system of linear equation whose vector form is $\sum_{i=1}^{n} \vec{e_i} x_i = \vec{0}$ or

$$\underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}}_{\vec{e}_1} x_1 + \underbrace{\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}}_{\vec{e}_2} x_2 + \cdots + \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}}_{\vec{e}_{n-1}} x_{n-1} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}_{\vec{e}_n} x_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

This system is in row-reduced Echelon form and has a unique solution

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

Thus the standard basis vectors are linearly independent. $\square$

Theorem 16. *Let $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ be collection of vectors. If $k = 1$ the system of vectors is linear dependent if and only if $\vec{v_1} = \vec{0}$.*

Proof. If $\vec{v} = \vec{0}$ then $1\vec{v} = 1\vec{0} = \vec{0}$ and therefore it is linearly dependent.

Assume $\vec{v}$ is linearly dependent then $\alpha\vec{v} = \vec{0}$ for some non-zero constant α. Multiplying both sides by α^{-1} we obtain $\vec{v} = \alpha^{-1}\vec{0}$. Or

$$\vec{v} = \alpha^{-1}\vec{0} = \alpha^{-1}(0\vec{0}) = \alpha^{-1}(0\vec{0}) = (\alpha^{-1}0)\vec{0} = (0)\vec{0} = 0\vec{0} = \vec{0}$$

$\square$

Theorem 17. *Let $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ be collection of vectors. If for some $1 \leq i \leq k$ we have that $\vec{v_i} = \vec{0}$ then the system of vectors is linear dependent.*

Proof. Self study excerise $\square$

Theorem 18. *Let $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ be collection of vectors. If for some $1 \leq i \neq j \leq k$ we have that $\vec{v_i} = \vec{v_j}$ then the system of vectors is linear dependent.*

Proof. Self study excerise $\square$

Theorem 19. *Let $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ be collection of linearly dependent vectors and $k > 1$. Then there is an index i such that $\vec{v_i}$ can be written as a linear combination of the remaining vectors.*

Proof. Self study excerise $\square$

2.6 Main theorem

The next result has various applications in determining if a set of vectors is linearly dependent as well as its size. Before stating it the following note is worth mentioning

Remark on terminology. Given two sets of vectors **A** and **B** such that every vector of **A** is a linear combination of the vectors in **B**, for example as in Theorem 12, consider the statement

A is a linear combination of vectors in **B**

and the statement

A is a linearly dependent set.

The former statement refers to Definition 34, the latter statement refers to Definition 35. To that end the statement

A is dependent

(bar any grammar issues) tends to be confusing. Does it mean **A** depends on how the vectors in **B** are given – that is if you change the vectors in **B**, the vectors in **A** also change? Or does it mean **A** is linearly dependent set? To my knowledge text on linear algebra would mean that **A** is linearly dependent in the sense of Definition 35. However, as far as students (or those who learn linear algebra) are concerned the statement meaning differs for different individuals. Please *do not* make such statements. Be verbose: use one of the previous two statements!

Theorem 20. *Let* $A = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_s}\}$ *and* $B = \left\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_k}\right\}$ *be two non-empty sets of vectors. Suppose that for each* $1 \leq i \leq s$ *we have that* $\vec{a_i}$ *is a linear combination of* $\left\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_k}\right\}$ *that is*

$$\begin{aligned}
\vec{a_1} &= \gamma_{11}\vec{b_1} + \gamma_{12}\vec{b_2} + \cdots + \gamma_{1k}\vec{b_k} \\
\vec{a_2} &= \gamma_{21}\vec{b_1} + \gamma_{22}\vec{b_2} + \cdots + \gamma_{2k}\vec{b_k} \\
&\vdots \\
\vec{a_s} &= \gamma_{s1}\vec{b_1} + \gamma_{s2}\vec{b_2} + \cdots + \gamma_{sk}\vec{b_k}.
\end{aligned}$$

Suppose also $s > k$ *then the vectors in* A *are a linearly dependent set of vectors.*

Proof. The argument proceeds by induction on k.

Base case $k = 1$: : Since $k = 1$ then $B = \{\vec{b_1}\}$. Then

$$\begin{aligned}
\vec{a_1} &= \gamma_{11}\vec{b_1} \\
\vec{a_2} &= \gamma_{21}\vec{b_1} \\
&\vdots \\
\vec{a_s} &= \gamma_{s1}\vec{b_1}.
\end{aligned}$$

If for any index i, $\gamma_{i1} = 0$ then A contains the zero vector and therefore A is linearly dependent. Suppose now for all indices i, $\gamma_{i1} \neq 0$ then since $s > k = 1$, there are at least two vectors $\vec{a_1}$ and $\vec{a_2}$ in A. Consider

$$\gamma_{21}\vec{a_1} - \gamma_{11}\vec{a_2} = \gamma_{21}\gamma_{11}\vec{b_1} - \gamma_{11}\gamma_{21}\vec{b_1} = 0\vec{b_1} = \vec{0}$$

Then $\vec{a_1}$ and $\vec{a_2}$ are linearly dependent. And since they are subset of A, then A itself is linearly dependent. This concludes the base case.

Inductive step: Let $k \geq 2$. By the theorem statement we have

$$\begin{aligned}
\vec{a_1} &= \gamma_{11}\vec{b_1} + \gamma_{12}\vec{b_2} + \cdots + \gamma_{1k}\vec{b_k} \\
\vec{a_2} &= \gamma_{21}\vec{b_1} + \gamma_{22}\vec{b_2} + \cdots + \gamma_{2k}\vec{b_k} \\
&\vdots \\
\vec{a_{s-1}} &= \gamma_{(s-1)1}\vec{b_1} + \gamma_{(s-1)2}\vec{b_2} + \cdots + \gamma_{(s-1)k}\vec{b_k}. \\
\vec{a_s} &= \gamma_{s1}\vec{b_1} + \gamma_{s2}\vec{b_2} + \cdots + \gamma_{sk}\vec{b_k}.
\end{aligned}$$

If all $\gamma_{s1}, \gamma_{s2}, \ldots, \gamma_{sk}$ are zero then $\vec{a_s} = \vec{0}$ and therefore A is linearly dependent. Suppose now at least one of $\gamma_{s1}, \gamma_{s2}, \ldots, \gamma_{sk}$ is non-zero. Without loss of generality let $\gamma_{sk} \neq 0$. In this case we add $\frac{-\gamma_{s1}}{\gamma_{sk}}$ the last equation to the first equation. Similarly, we add $\frac{-\gamma_{s2}}{\gamma_{sk}}$ the last equation to the second equation and so forth until we add $\frac{-\gamma_{(s-1)k}}{\gamma_{sk}}$ the last equation to equation $s-1$ to obtain equations

$$\begin{aligned}
\vec{a'_1} = \vec{a_1} - \frac{\gamma_{1k}}{\gamma_{sk}}\vec{a_s} &= \gamma'_{11}\vec{b_1} + \gamma'_{12}\vec{b_2} + \cdots + \gamma'_{1(k-1)}\vec{b}_{(k-1)} \\
\vec{a'_2} = \vec{a_2} - \frac{\gamma_{2k}}{\gamma_{sk}}\vec{a_s} &= \gamma'_{21}\vec{b_1} + \gamma'_{22}\vec{b_2} + \cdots + \gamma'_{2(k-1)}\vec{b}_{(k-1)} \\
&\vdots \\
\vec{a'_{s-1}} = \vec{a_{s-1}} - \frac{\gamma_{(s-1)k}}{\gamma_{sk}}\vec{a_s} &= \gamma'_{(s-1)1}\vec{b_1} + \gamma'_{(s-1)2}\vec{b_2} + \cdots + \gamma'_{(s-1)(k-1)}\vec{b}_{k-1}.
\end{aligned}$$

Since $s > k$, then $s - 1 > k - 1$. Furthermore, each $\vec{a'_1}, \ldots, \vec{a'_{s-1}}$ is a linear combination of the vectors $\vec{b'_1}, \ldots, \vec{b'_{k-1}}$. We apply the inductive hypothesis to conclude that $\vec{a'_1}, \ldots, \vec{a'_{s-1}}$ are linearly dependent. In other words there exists $\mu_1, \mu_2, \ldots, \mu_{s-1}$ not all zero such that

$$\begin{aligned}
\vec{0} &= \mu_1\vec{a'_1} + \mu_2\vec{a'_2} + \cdots + \mu_{s-1}\vec{a'_{s-1}} \\
&= \mu_1\left(\vec{a_1} - \frac{\gamma_{1k}}{\gamma_{sk}}\vec{a_s}\right) + \mu_2\left(\vec{a_2} - \frac{\gamma_{2k}}{\gamma_{sk}}\vec{a_s}\right) + \mu_{s-1}\left(\vec{a_{s-1}} - \frac{\gamma_{(s-1)k}}{\gamma_{sk}}\vec{a_s}\right) \\
&= \mu_1\vec{a_1} + \mu_2\vec{a_2} + \cdots + \mu_{s-1}\vec{a_{s-1}} + \tau\vec{a_s}.
\end{aligned}$$

Since at least one of μ_i's is non-zero, the vectors $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_s}$ are linearly dependent. $\square$

Example: In the standard vector space $\mathbb{R}^4$ if a set $\mathbf{A}$ contains five or more vectors $\vec{a_i} = \begin{pmatrix} a_{i1} \\ a_{i2} \\ a_{i3} \\ a_{i4} \end{pmatrix}$ each one is a linear combination of the standard basis vectors

$$\vec{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \vec{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \vec{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \vec{e_4} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

namely,

$$\vec{a_i} \quad = \quad a_{i1}\vec{e}_1 + a_{i2}\vec{e}_2 + a_{i3}\vec{e}_3 + a_{i4}\vec{e}_4.$$

Then the vectors in $\mathbf{A}$ are linearly dependent. This can be generalized to say that in $\mathbb{R}^n$ any set of $n+1$ vectors is linearly dependent. Observe that if we have n or less vectors in $\mathbb{R}^n$ they may or may not be linearly independent. Theorem 20 implies nothing if the size of $\mathbf{A}$ is smaller than or equal to the size of $\mathbf{B}$.

Example: Let $\mathbf{A} = \{\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4\}$ and $\mathbf{B} = \left\{\vec{b}_1, \vec{b}_2, \vec{b}_3\right\}$ where

$$\vec{a_1} = \begin{pmatrix} 3 \\ -3 \\ 4 \\ 7 \end{pmatrix} \quad \vec{a}_2 = \begin{pmatrix} 6 \\ 0 \\ 4 \\ 10 \end{pmatrix} \quad \vec{a}_3 = \begin{pmatrix} 3 \\ 0 \\ 2 \\ 5 \end{pmatrix} \quad \vec{a}_4 = \begin{pmatrix} 4 \\ 7 \\ 0 \\ 2 \end{pmatrix}$$

and

$$\vec{b}_1 = \begin{pmatrix} 2 \\ -1 \\ 2 \\ 4 \end{pmatrix} \quad \vec{b}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \quad \vec{b}_3 = \begin{pmatrix} -2 \\ -5 \\ 0 \\ 0 \end{pmatrix}$$

We have that

$$\begin{aligned} \vec{a}_1 &= 2\vec{b}_1 - \vec{b}_2 \\ \vec{a}_2 &= 2\vec{b}_1 + 2\vec{b}_2 \\ \vec{a}_3 &= \vec{b}_1 + \vec{b}_2 \\ \vec{a}_4 &= 2\vec{b}_2 - \vec{b}_3 \end{aligned}$$

then vectors in $\mathbf{A}$ are linearly dependent.

Remark: in the first example above the set $\mathbf{B}$ is linearly independent (the vectors $\vec{e}_i$ are linearly independent). In the second example the set of vectors in $\mathbf{B}$ are linearly dependent. This shows that Theorem 20 says nothing about the linear dependence or independence of the set $\mathbf{B}$.

The next result show a non-trivial application of Theorem 20.

Theorem 21. *An $n \times n$ matrix A has linearly dependent rows if and only if it has linearly dependent columns.*

Proof. Let $A = (\vec{c}_1, \ldots, \vec{c}_n) = \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_n \end{pmatrix}$. If the rows of A are linearly dependent then we know that the equation

$$\alpha_1 \vec{r}_1 + \cdots + \alpha_n \vec{r}_n = \vec{0}$$

has a solution where at least one $\alpha_i \neq 0$. Then we can write

$$\vec{r}_i = \frac{\alpha_1}{\alpha_i}\vec{r}_1 + \cdots + \frac{\alpha_{i-1}}{\alpha_i}\vec{r}_{i-1} + \frac{\alpha_1}{\alpha_{i+1}}\vec{r}_{i+1} + \cdots + \frac{\alpha_n}{\alpha_i}\vec{r}_n$$

Thus each vector in the set $\{\vec{r}_1, \ldots, \vec{r}_n\}$ is a linear combination of the vectors in $\{\vec{r}_1, \ldots, \vec{r}_{i-1}, \vec{r}_{i+1}, \ldots, \vec{r}_n\}$. Therefore using properties of matrix multiplication we can write

$$A = \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_n \end{pmatrix} = BC$$

Where B is an $n \times n-1$ matrix and C is a $n-1 \times n$ matrix. Furthermore the rows of C are equal to $\{\vec{r}_1, \ldots, \vec{r}_{i-1}, \vec{r}_{i+1}, \ldots, \vec{r}_n\}$. In this case, however, the columns of A are linear combinations of the columns of B. There are $n-1$ columns in B and n columns in A. Thus by the above Theorem 20 the columns of A are linearly dependent. Similar argument applies if the columns of A are linearly dependent. ☐

2.7 Subspaces

Definition 36 (subspace)**.** *Let $\mathbf{V}$ be a vector space and let $\mathbf{U}$ be a subset of $\mathbf{V}$. If $\mathbf{U}$ is a vector space itself then $\mathbf{U}$ is called a* subspace *of $\mathbf{V}$*

Examples: the following are example of subspaces.

1. Every vector space is a subspace of itself.

2. The x-axis defined as
$$\left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}$$
is subspace of $\mathbb{R}^2$

3. The y-axis defined as
$$\left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \mid y \in \mathbb{R} \right\}$$
is subspace of $\mathbb{R}^2$

4. The set
$$\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$
is a subspace of $\mathbb{R}^2$.

5. $\mathbf{P}_n$ polynomials of degree at most n are a subspace of the set of all polynomials $\mathbf{P}$.

6. For the vector space **CVS** described in §2.2 the vectors the following set is a subspace
$$\left\{ \begin{bmatrix} 2 \\ y \end{bmatrix} \mid y \in \mathbb{R} \right\}$$

7. For the vector space **CVS** described in §2.2 the vectors the following set is a subspace
$$\left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \mid x \in \mathbb{R} \right\}$$

8. For the vector space **CVS** described in §2.2 the vectors the following set is a subspace
$$\left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\}$$
The set consists just of the zero vector in **CVS**.

Counterexamples: the following are not subspaces (even though they are subsets of the corresponding vector spaces).

1. line through points $(1, 0)$ and $(0, 1)$
$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid x + y = 1 \right\}$$
is *not* a subspace of $\mathbb{R}^2$. It does not contain the zero vector for example; the vector operations are also not closed.

2. The set consisting of the single element
$$\left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\}$$
is *not* a subspace of $\mathbb{R}^2$.

3. Even degree polynomials $\mathbf{P}_e$ are *not* a subspace of polynomials $\mathbf{P}$: for example
$$\begin{aligned} p_1(x) &= x^2 + x + 1 \in \mathbf{P}_e \\ p_2(x) &= -x^2 - 1 \in \mathbf{P}_e \end{aligned}$$

but

$$p_1(x) + p_2(x) = x \notin \mathbf{P}_e$$

thus polynomial addition is not closed for even degree polynomials.

4. Odd degree polynomials $\mathbf{P}_o$ are *not* a subspace of polynomials $\mathbf{P}$: for example

$$\begin{aligned} q_1(x) &= x^3 + x^2 + 1 \in \mathbf{P}_o \\ q_2(x) &= -x^3 - 1 \in \mathbf{P}_o \end{aligned}$$

but

$$q_1(x) + q_2(x) = x^2 \notin \mathbf{P}_o$$

thus polynomial addition is not closed for odd degree polynomials.

5. For the vector space **CVS** described in §2.2 the set

$$\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

is not a subspace of **CVS**.

Theorem 22. *In any vector space* $\mathbf{V}$ *the vector space itself is a subspace and the zero vector on its own is a vector space.*

Proof. Verification of all vector space properties is straightforward. ☐

The vector space itself and the zero vector are often called the *trivial* subspaces. The next result establishes an efficient way to test of a subset of a vectors space is also as subspace.

Theorem 23. *A set* $\mathbf{U}$ *is a subspace of* $\mathbf{V}$ *if and only if for all* $\vec{u}, \vec{w} \in \mathbf{U}$ *and for all* $s, t \in \mathbb{K}$ *we have that* $s\vec{u} + t\vec{w} \in \mathbf{U}$.

Proof. If $\mathbf{U}$ is a subspace of a vector space $\mathbf{V}$, then it is a vector space itself and since the operations are closed we have that $s\vec{u} + t\vec{w} \in \mathbf{U}$.

Suppose now $\forall \vec{u}, \vec{w} \in \mathbf{U}, \forall s, t \in \mathbb{K}, s\vec{u} + t\vec{v} \in \mathbf{U}$. We will verify all properties of vector spaces for $\mathbf{U}$.

closure of + from $\forall \vec{u}, \vec{w} \in \mathbf{U}, \forall s, t \in \mathbb{K}, s\vec{u} + t\vec{w} \in \mathbf{U}$ for $s = t = 1$ we conclude $\forall \vec{u}, \vec{w} \in \mathbf{U}, \vec{u} + \vec{w} = 1\vec{u} + 1\vec{w} \in \mathbf{U}$

closure of $\cdot$ from $\forall \vec{u}, \vec{w} \in \mathbf{U}, \forall s, t \in \mathbb{K}, s\vec{u} + t\vec{w} \in \mathbf{U}$ for $t = 0$ and $\vec{u} = \vec{w}$ we conclude $\forall \vec{u} \in \mathbf{U}, \forall s \in \mathbb{K}, s\vec{u} = s\vec{u} + 0\vec{w} \in \mathbf{U}$

commutativity since the operations are inherited from $\mathbf{V}$ the result follows

associativity since the operations are inherited from $\mathbf{V}$ the result follows

zero vector from $\forall \vec{u}, \vec{w} \in \mathbf{U}, \forall s, t \in \mathbb{K}, s\vec{u} + t\vec{w} \in \mathbf{U}$ for $s = t = 0$ $\vec{0} = \vec{0} + \vec{0} = 0\vec{u} + 0\vec{w} \in \mathbf{U}$; uniqueness and neutrality is inherited from $\mathbf{V}$;

neutral inverse from $\forall \vec{u}, \vec{w} \in \mathbf{U}, \forall s, t \in \mathbb{K}, s\vec{u} + t\vec{w} \in \mathbf{U}$ for $\vec{u} = \vec{w}$, $s = 0$ and $t = -1$ we have $\forall \vec{u} \in \mathbf{U}, \vec{-u} = \vec{0} + \vec{-u} = 0\vec{u} + (-1)\vec{u} \in \mathbf{U}$;

distributive properties are inherited from $\mathbf{V}$.

$\square$

Example: consider

$$\mathbb{S} = \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mid 3X + 6Y = 2Z \right\} \subset \mathbb{R}^3.$$

One way to verify the above subset is also a subspace is to check all conditions of Definition 33 as done in §2.2. The alternative is to use Theorem 23. Let

$$\begin{aligned} \vec{u} &= \begin{pmatrix} x_u \\ y_u \\ z_u \end{pmatrix} \in \mathbb{S} \\ \vec{w} &= \begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} \in \mathbb{S} \end{aligned}$$

By definition we have

$$\begin{aligned} 3x_u + 6y_u &= 2z_u \\ 3x_w + 6y_w &= 2z_w. \end{aligned}$$

Multiply the first equation with s and the second equation with t and add them together to obtain

$$3(sx_u + tx_w) + 6(sx_u + ty_w) = 2(sz_u + tz_w)$$

which means

$$\begin{aligned} s\vec{u} + t\vec{w} &= s\begin{pmatrix} x_u \\ y_u \\ z_u \end{pmatrix} + t\begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} \\ &= \begin{pmatrix} 3(sx_u + tx_w) \\ 6(sx_u + ty_w) \\ 2(sz_u + tz_w) \end{pmatrix} \in \mathbb{S} \end{aligned}$$

By Theorem 23 the set $\mathbb{S}$ is a subspace of $\mathbb{R}^3$.

Example: consider

$$\mathbb{T} = \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mid 3X + Y = Z + 1 \right\} \subset \mathbb{R}^3.$$

Let

$$\begin{aligned}\vec{u} &= \begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix} \in \mathbb{T} \\ \vec{w} &= \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \in \mathbb{T}\end{aligned}$$

with $s = 1$ and $t = 1$ we get

$$\begin{aligned}s\vec{u} + t\vec{w} &= 1\begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 2 \\ 6 \end{pmatrix} \notin \mathbb{T}\end{aligned}$$

By Theorem 26 the set $\mathbb{T}$ is *not* a subspace of $\mathbb{R}^3$.

2.8 Span

Definition 37 (span). *Let $\mathbf{S} = \{\vec{u}_1, \ldots, \vec{u}_m\}$ be a set of vector the set of all linear combinations of the vectors in $\mathbf{S}$ is called the* span *of $\mathbf{S}$ and denoted by $\langle \mathbf{S} \rangle$*

$$\langle \mathbf{S} \rangle = \{a_1\vec{u}_1 + \cdots + a_m\vec{u}_m \mid a_1 \ldots a_m \in \mathbb{K}\}$$

If $\mathbf{S} = \emptyset$ then $\langle \mathbf{S} \rangle = \{\vec{0}\}$

Examples:

1. For the vector space $\mathbb{R}^2$ we have

$$\begin{aligned}\mathbb{R}^2 &= \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\rangle \\ x\text{-axis} &= \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \\ y\text{-axis} &= \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\rangle \\ \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} &= \left\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\rangle = \langle \emptyset \rangle\end{aligned}$$

If l is the line $x+y=0$ in $\mathbb{R}^2$ that is the set

$$l = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid x+y=0 \right\}$$

then

$$l = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle$$

2. $\langle 1, x, x^2 \rangle = \mathbf{P}_2$ that is the span of x^0, x and x^2 is the set of all polynomials of degree at most two.

3. $\langle \mathbf{P}_2 \rangle = \mathbf{P}_2$ that is the span of all polynomials of degree at most two is the set of all polynomials of degree at most two.

4. $\langle 1, x, x^2, \ldots, x^n \rangle = \mathbf{P}_n$ that is the span of $x^0, x, \ldots, x^n$ is the set of all polynomials of degree at most n.

5. $\langle x^0, x^1, x^2, \ldots, x^n \ldots \rangle = \mathbf{P}$ that is the span of powers of x is the set of all polynomials.

6. $\langle \mathbf{P} \rangle = \mathbf{P}$ that is the span of polynomials is the set of all polynomials.

7. For any vector space $\mathbf{V}$ we have

$$\langle \mathbf{V} \rangle = \mathbf{V}$$

8. For the vector space **CVS** described in §2.2 we have

$$\begin{aligned}
\mathbf{CVS} &= \left\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\rangle \\
\left\{ \begin{bmatrix} 2 \\ y \end{bmatrix} \mid y \in \mathbb{R} \right\} &= \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle \\
\left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \mid x \in \mathbb{R} \right\} &= \left\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\rangle \\
\left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\} &= \left\langle \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\rangle = \langle \emptyset \rangle
\end{aligned}$$

9. Let $\vec{s}_t$ denote the sequence $\{s_{ti}\}$ where

$$s_{ti} = \begin{cases} 1 & t = i \\ 0 & t \neq i \end{cases}$$

that is

$$\begin{aligned}
\vec{s}_1 &= 1, 0, 0, 0, \ldots \\
\vec{s}_2 &= 0, 1, 0, 0, \ldots \\
\vec{s}_3 &= 0, 0, 1, 0, \ldots \\
&\vdots
\end{aligned}$$

Then the set of sequences with finitely may non-zero terms equals

$$\langle \vec{s}_1, \vec{s}_2, \dots \rangle$$

Example: consider the system of linear equations

$$\begin{aligned} x_1 + 2x_2 + x_3 + 4x_4 + 2x_5 &= 5 \\ 2x_1 + 4x_2 + 3x_3 + 8x_4 + 3x_5 &= 9 \\ x_1 + 2x_2 + 2x_3 + 5x_4 + x_5 &= 5 \\ x_1 + 2x_2 + x_3 + 2x_5 &= 1 \end{aligned}$$

equivalently in vector form

$$\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix} x_2 + \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix} x_3 + \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix} x_4 + \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} x_5 \quad = \quad \begin{pmatrix} 5 \\ 9 \\ 5 \\ 1 \end{pmatrix}$$

has solution (verify it)

$$\begin{aligned} x_1 &= 3 + 3s_1 + 2s_2 \\ x_2 &= 1 + s_2 \\ x_3 &= -2 - s_1 \\ x_4 &= 1 \\ x_5 &= -1 + s_1 \end{aligned}$$

in vector form

$$\left\{ \begin{pmatrix} 3 \\ 1 \\ -2 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \\ -1 \\ 0 \\ -1 \end{pmatrix} s_1 + \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} s_2 \mid s_1, s_2 \in \mathbb{R} \right\}.$$

The particular solution implies

$$3 \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \quad = \quad \begin{pmatrix} 5 \\ 9 \\ 5 \\ 1 \end{pmatrix}$$

equivalently

$$\begin{pmatrix} 5 \\ 9 \\ 5 \\ 1 \end{pmatrix} \in \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \right\rangle.$$

Conversely, if

$$\begin{pmatrix} 4 \\ 9 \\ 5 \\ 4 \end{pmatrix} \in \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \right\rangle$$

then the system of linear equations in vector form

$$\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix} x_2 + \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix} x_3 + \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix} x_4 + \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} x_5 \quad = \quad \begin{pmatrix} 4 \\ 9 \\ 5 \\ 4 \end{pmatrix}$$

has a solution. This is an example of the following result

Theorem 24. *A system $Ax = b$ has a solution if and only if b is in the span of the columns of A.*

Proof. Consider $A\vec{x} = \vec{b}$ as matrix multiplication by Theorem 13 the columns of the result are a linear combination of the columns of A. Thus if $A\vec{x} = \vec{b}$ has a solution then $\vec{b}$ is a linear combination of the columns of A equivalently $\vec{b}$ is in the span of the columns of matrix A. Conversely, if $\vec{b}$ is in the span of the columns of matrix A then there is (column) vector $\vec{c}$ such that $A\vec{c} = \vec{b}$ and therefore $A\vec{x} = \vec{b}$ has a solution. $\square$

Example: continuing the above example using the homogeneous solutions we have

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

For the set $\langle A \rangle$ which is the span of columns of matrix A by Theorem 12 we get

$$\begin{aligned}
\langle A \rangle &= \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \right\rangle \\
&= \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \right\rangle \\
&= \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix} \right\rangle
\end{aligned}$$

Theorem 25. *$\langle \mathbf{S} \rangle = \langle \mathbf{S} \cup \vec{u} \rangle$ if and only if $\vec{u} \in \langle \mathbf{S} \rangle$.*

Proof. Assume first $\vec{u} \in \langle \mathbf{S} \rangle$. The inclusion $\langle \mathbf{S} \rangle \subseteq \langle \mathbf{S} \cup \vec{u} \rangle$ holds since $\mathbf{S} \subseteq \mathbf{S} \cup \vec{u}$. For $\langle \mathbf{S} \rangle \supseteq \langle \mathbf{S} \cup \vec{u} \rangle$ since $\vec{u} \in \langle \mathbf{S} \rangle$ then $\vec{u}$ is a linear combination of the vectors in $\mathbf{S}$ and by Theorem 12 any vector that is linear combination of $\mathbf{S} \cup \vec{u}$ is a linear combination of the vectors in $\mathbf{S}$ meaning that $\langle \mathbf{S} \rangle \supseteq \langle \mathbf{S} \cup \vec{u} \rangle$.

Assume now $\langle \mathbf{S} \cup \vec{u} \rangle = \langle \mathbf{S} \rangle$ since $\vec{u} \in \mathbf{S} \cup \vec{u}$ then $\vec{u} \in \langle \mathbf{S} \cup \vec{u} \rangle = \langle \mathbf{S} \rangle$.

Thus the result follows. □

Theorem 26. *The span of a set of vectors is a vector space.*

Proof. Let $\vec{u}, \vec{w} \in \langle \mathbf{S} \rangle$, then by properties of vector spaces $s\vec{u} + t\vec{w} \in \langle \mathbf{S} \rangle$. By Theorem 23 the result follows. □

2.9 Basis and Dimension

Definition 38 (basis)**.** *Let $\mathbf{V}$ be a vector space, the set of vectors $\mathbf{B} = \{\vec{b}_1, \ldots, \vec{b}_d\}$ is a* basis *for $\mathbf{V}$ if every vector in $\mathbf{V}$ can be represented as a linear combination of the vectors in $\mathbf{B}$ and the vectors in $\mathbf{B}$ are linearly independent.*

We can express elements of a vector space using basis elements. If $\mathbf{B}$ is a basis for $\mathbf{V}$ then $\mathbf{B} \subseteq \mathbf{V}$.

Standard basis: For the vector space $\mathbb{R}^3$ the set $\mathbf{E} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ where

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

is a basis since for any vector $\vec{u} \in \mathbb{R}^3$ we have $\vec{u} \in \langle E \rangle$. This basis is known as the standard basis and can be generalized to any $\mathbb{R}^n$. An alternative basis for $\mathbb{R}^3$ is $\mathbf{B} = \left\{ \vec{b}_1, \vec{b}_2, \vec{b}_3 \right\}$ where

$$\vec{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \vec{b}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \vec{b}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Using span notation we have

$$\mathbb{R}^3 = \langle \vec{e}_1, \vec{e}_2, \vec{e}_3 \rangle = \left\langle \vec{b}_1, \vec{b}_2, \vec{b}_3 \right\rangle$$

Example: for the subspace $\mathbb{S}$ of the vector space $\mathbb{R}^3$ defined in §2.7 we have

$$\begin{aligned} \mathbb{S} = \langle \mathbf{B} \rangle &= \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \right\rangle \\ \langle \mathbf{D} \rangle &= \left\langle \begin{pmatrix} 4 \\ -5 \\ -9 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix} \right\rangle \end{aligned}$$

Example: for the vector space $\mathbf{P}_2$ we have bases $\mathbf{B}$ and $\mathbf{D}$ where

$$\begin{aligned} \mathbf{P}_2 = \langle \mathbf{B} \rangle &= \left\langle 6\,x^2 - 39\,x + 23, \quad -5\,x^2 + 33\,x - 19, \quad -x^2 + 7\,x - 4 \right\rangle \\ \langle \mathbf{D} \rangle &= \left\langle x^2 - 8\,x + 5, \quad x^2 - 5\,x + 3, \quad -x^2 + 7\,x - 4 \right\rangle \end{aligned}$$

Example: consider the vector space **CVS** described in §2.2. It has bases $\mathbf{B}$ and $\mathbf{D}$ where

$$\begin{aligned} \mathbf{CVS} = \langle \mathbf{B} \rangle &= \left\langle \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\rangle \\ \langle \mathbf{D} \rangle &= \left\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\rangle \end{aligned}$$

Remark: in all examples above the number of basis vector for the same vectors space remains the same. This is not a coincidence.

Theorem 27. *Let $\mathbf{B}_1$ and $\mathbf{B}_2$ be two distinct basis for a vector space $\mathbf{V}$. Then the number of vectors in $\mathbf{B}_1$ and $\mathbf{B}_2$ is the same.*

Proof. Suppose by contradiction the size (number of vectors in the set) of $\mathbf{B}_1$ does not equal the size of $\mathbf{B}_2$. Without loss of generality suppose the size $\mathbf{B}_2$ is strictly larger than the size of $\mathbf{B}_1$. By definition of basis every vector in $\mathbf{B}_2$ is a linear combination of the vectors in $\mathbf{B}_1$. Then by Theorem 20 the vectors in $\mathbf{B}_2$ must be linearly dependent contradicting the fact that $\mathbf{B}_2$ is a basis. Thus all basis have the same number of vectors. $\square$

Definition 39 (dimension)**.** *Let* $\mathbf{B}$ *be a basis for a vector space* $\mathbf{V}$*, then the size of* $\mathbf{B}$ *is called the* dimension *of* $\mathbf{V}$*.*

Definition 40 (finite dimensional vector space)**.** *A vector space is called finite dimensional if it has a basis with only finitely many vectors.*

Remark many of the result that are discussed here are valid for infinite dimensional vector space, however, the arguments presented here are valid for finite dimensional vector space. Infinite dimensional vector spaces are a topic of a different course. Using the notation/examples from §2.1

Examples:

1. $\mathbb{K}^n$ is finite dimensional vector space;
2. $\mathcal{M}_{n\times m}(\mathbb{K})$ is finite dimensional vector space;
3. $\mathcal{C}[a, b]$ is infinite dimensional vector space;
4. $\mathbf{P}$ is infinite dimensional vector space;
5. $\mathbf{P}_n$ is finite dimensional vector space;
6. the vector space from §2.2 is finite dimensional vector space.

Theorem 28. *Any linearly independent set can be extended to a basis.*

Proof. Self study exercise. $\square$

Theorem 29. *Any spanning set contains a basis.*

Proof. Self study exercise. $\square$

2.10 Coordinates

In § 2.4 the underlying question is whether a system of linear equations has a solution. Span is the set of all linear combination so §2.8 reiterates the same problem. Linear independence as in §2.5 is concerned with uniqueness of solution to a homogeneous system of linear equations. For a given consistent system of linear equations if the corresponding homogeneous system of linear equations has a unique solution then the original system of linear equations has a unique solution. Coordinates, similar to linear independence, discuss the idea of non-homogeneous system of linear equations having a unique solution.

Theorem 30. *Let $\vec{e_1}, \ldots, \vec{e_n}$ be linearly independent. Suppose $\vec{u} = a_1\vec{e_1} + \cdots + a_n\vec{e_n}$ and $\vec{u} = b_1\vec{e_1} + \cdots + b_n\vec{e_n}$. Then $a_1 = b_1$, $a_2 = b_2$,..., $a_n = b_n$.*

Proof. Suppose

$$\begin{aligned} \vec{u} &= a_1\vec{e_1} + \cdots + a_n\vec{e_n} \\ \vec{u} &= b_1\vec{e_1} + \cdots + b_n\vec{e_n}. \end{aligned}$$

Subtracting them gives

$$\vec{0} = (a_1 - b_1)\vec{e_1} + \cdots + (a_n - b_n)\vec{e_n}.$$

Since $\vec{e_1}, \ldots, \vec{e_n}$ are linearly independent the only solution is

$$\begin{aligned} 0 &= a_1 - b_1 \\ &\vdots \\ 0 &= a_n - b_n \end{aligned}$$

which implies the result. $\square$

Definition 41 (coordinates)**.** *Let $\vec{u}$ be a vector in a d-dimensional vector space $\mathbf{V}$. Let $\mathbf{B} = \{\vec{b}_1, \ldots, \vec{b}_d\}$ be a basis for $\mathbf{V}$. By definition of basis*

$$\vec{u} = u_1\vec{b}_1 + \cdots + u_d\vec{b}_d.$$

The values $u_1, \ldots, u_d$ are the coordinates of $\vec{u}$ with respect to basis $\mathbf{B}$.

Representation map. Let $\mathbf{U}$ be a vector space of dimension d with a basis $\mathbf{B} = \{\vec{b}_1, \ldots, \vec{b}_d\}$. The map

$$\mathcal{R}_{\mathbf{B}} : \mathbf{U} \to \mathbb{K}^d$$

called *representation map* is defined as

$$\mathcal{R}_{\mathbf{B}}(\vec{u}) = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{pmatrix}_{\mathbf{B}}$$

where $u_1, u_2, \ldots, u_d$ are the coordinates of $\vec{u}$ in basis $\mathbf{B}$. The vector $\begin{pmatrix} u_1 \\ \vdots \\ u_d \end{pmatrix}_{\mathbf{B}}$ is the representation of vector $\vec{u}$ in basis $\mathbf{B}$. Whenever the basis $\mathbf{B}$ is understood from the context, the subscript is often omitted. It is straightforward to verify that for a basis $\mathbf{B}$ if

$$\mathcal{R}_{\mathbf{B}}(\vec{u}) = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{pmatrix}_{\mathbf{B}}$$

and

$$\mathcal{R}_{\mathbf{B}}\left(\vec{w}\right)=\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{pmatrix}_{\mathbf{B}}$$

then

$$\mathcal{R}_{\mathbf{B}}\left(s\vec{u}+t\vec{w}\right)=s\mathcal{R}_{\mathbf{B}}\left(\vec{u}\right)+\mathcal{R}_{\mathbf{B}}\left(t\vec{w}\right)=\begin{pmatrix} su_1+tw_1 \\ su_2+tw_2 \\ \vdots \\ su_3+tw_d \end{pmatrix}_{\mathbf{B}}$$

Example: consider the vectors space $\mathbb{R}^3$. For vector $\vec{u}$ where

$$\vec{u}=\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

and the standard basis $\mathbf{E}=\{\vec{e}_1,\vec{e}_2,\vec{e}_3\}$ where

$$\vec{e}_1=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\vec{e}_2=\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\vec{e}_3=\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

the vector equation

$$\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}+x_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}+x_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

is equivalent to a system of linear equations with augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

whose solution particular solution are the coordinates of $\vec{u}$ in basis $\mathbf{E}$ and implies the vector equality

$$\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}+0\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}+2\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

In other words the representation of $\vec{u}$ in basis $\mathbf{E}$ is

$$\mathcal{R}_E\left(\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}_E$$

For the same vector space and a different basis $\mathbf{B} = \left\{\vec{b}_1, \vec{b}_2, \vec{b}_3\right\}$ where

$$\vec{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \vec{b}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \vec{b}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

the vector equation

$$\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

corresponds to a system of linear equations with augmented matrix

$$\left(\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 2 \end{array}\right)$$

whose particular solution consists of the coordinates of vector $\vec{u}$ in basis $\mathbf{B}$ and implies the vector equality

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

In other words the representation of $\vec{u}$ in basis $\mathbf{B}$ is

$$\mathcal{R}_B\left(\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}_B$$

Example: in the vectors space $\mathbb{R}^3$ the subset

$$\mathbb{S} = \left\{\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mid 3X + 6Y = 2Z\right\}$$

is a subspace, which means it is a vector space on its own. For vector $\vec{u} \in \mathbb{S}$ where

$$\vec{u} = \begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix}$$

and basis $\mathbf{B} = \left\{\vec{b}_1, \vec{b}_2\right\}$ where

$$\vec{b}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \vec{b}_2 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$$

the vector equation

$$\begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix} = x_1 \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$$

is equivalent to a system of linear equations with augmented matrix

$$\left(\begin{array}{cc|c} 2 & 0 & 4 \\ -1 & 1 & -3 \\ 0 & 3 & -3 \end{array}\right)$$

whose particular solutions are the coordinates of $\vec{u}$ in basis $\mathbf{B}$ and imply the vector equality

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

In other words the representation of vector $\vec{u}$ in basis $\mathbf{B}$ is

$$\mathcal{R}_B\left(\begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}_B$$

Consider an alternative basis $\mathbf{D} = \left\{\vec{d_1}, \vec{d_2}\right\}$ where

$$\vec{d_1} = \begin{pmatrix} 4 \\ -5 \\ -9 \end{pmatrix}, \vec{d_2} = \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix}$$

for the same vector $\vec{u}$ the vector equation

$$\begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix} = x_1 \begin{pmatrix} 4 \\ -5 \\ -9 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix}$$

is equivalent to a system of linear equations with augmented matrix

$$\left(\begin{array}{cc|c} 4 & 2 & 4 \\ -5 & -3 & -3 \\ -9 & -6 & -3 \end{array}\right)$$

whose particular solutions are the coordinates of $\vec{u}$ in basis $\mathbf{D}$ and imply the vector equality

$$\begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix} = 3 \begin{pmatrix} 4 \\ -5 \\ -9 \end{pmatrix} - 4 \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix}$$

In other words the representation of vector $\vec{u}$ in basis $\mathbf{D}$ is

$$\mathcal{R}_D\left(\begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix}\right) = \begin{pmatrix} 3 \\ -4 \end{pmatrix}_D$$

Example: consider $\mathbf{P}_2$ – the vector space of polynomials of degree at most two. For vector $\vec{u}$ where

$$\vec{u} = -x^2 + 5\,x + 3$$

and basis $\mathbf{B} = \left\{\vec{b}_1, \vec{b}_2, \vec{b}_3\right\}$ where

$$\begin{aligned} \vec{b}_1 &= 6\,x^2 - 39\,x + 23 \\ \vec{b}_2 &= -5\,x^2 + 33\,x - 19 \\ \vec{b}_3 &= -x^2 + 7\,x - 4. \end{aligned}$$

For the same vector $\vec{u}$ the vector equation

$$-x^2 + 5\,x - 3 \quad = \quad \alpha_1\left(6\,x^2 - 39\,x + 23\right) + \alpha_2\left(-5\,x^2 + 33\,x - 19\right) + \alpha_3\left(-x^2 + 7\,x - 4\right)$$

is equivalent to a system of linear equations with augmented matrix

$$\left(\begin{array}{rrr|r} 23 & -19 & -4 & -3 \\ -39 & 33 & 7 & 5 \\ 6 & -5 & -1 & -1 \end{array}\right)$$

whose particular solution are the coordinates of vector $\vec{u}$ in basis $\mathbf{B}$ and imply the vector equality

$$-x^2 + 5\,x - 3 \quad = \quad 0\left(6\,x^2 - 39\,x + 23\right) + 1\left(-5\,x^2 + 33\,x - 19\right) - 4\left(-x^2 + 7\,x - 4\right)$$

In other words the representation of vector $\vec{u}$ in basis $\mathbf{B}$ is

$$\mathcal{R}_B\left(-x^2 + 5\,x - 3\right) \quad = \quad \left(\begin{array}{r} 0 \\ 1 \\ -4 \end{array}\right)_B .$$

Consider a different basis $\mathbf{D} = \left\{\vec{d}_1, \vec{d}_2, \vec{d}_3\right\}$ where

$$\begin{aligned} \vec{d}_1 &= x^2 - 8\,x + 5 \\ \vec{d}_2 &= x^2 - 5\,x + 3 \\ \vec{d}_3 &= -x^2 + 7\,x - 4 \end{aligned}$$

the vector equation

$$-x^2 + 5\,x - 3 \quad = \quad \alpha_1\left(x^2 - 8\,x + 5\right) + \alpha_2\left(x^2 - 5\,x + 3\right) + \alpha_3\left(-x^2 + 7\,x - 4\right)$$

is equivalent to a system of linear equations with augmented matrix

$$\left(\begin{array}{rrr|r} 5 & 3 & -4 & -3 \\ -8 & -5 & 7 & 5 \\ 1 & 1 & -1 & -1 \end{array}\right)$$

whose particular solution are the coordinates of vector $\vec{u}$ in basis $\mathbf{D}$ and imply the vector equality

$$-x^2 + 5\,x - 3 \quad = \quad 0\left(x^2 - 8\,x + 5\right) - 1\left(x^2 - 5\,x + 3\right) + 0\left(-x^2 + 7\,x - 4\right)$$

. In other words the representation of vector $\vec{u}$ in basis $\mathbf{D}$ is

$$\mathcal{R}_D\left(-x^2 + 5\,x - 3\right) \quad = \quad \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}_D .$$

Example: consider the vector space **CVS** described in §2.2. For vector $\vec{u}$ where

$$\vec{u} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

and basis $B = \left\{\vec{b}_1, \vec{b}_2\right\}$ where

$$\vec{b}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

By equating the components in the vector equation

$$\begin{aligned} \begin{bmatrix} 4 \\ -3 \end{bmatrix} &= x_1 \odot \begin{bmatrix} 3 \\ 0 \end{bmatrix} \oplus x_2 \odot \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} x_1 + 2 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 2 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1 + 2 \\ x_2 \end{bmatrix} \end{aligned}$$

one obtains a system of linear equation with augmented matrix

$$\left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -3 \end{array}\right).$$

whose solution implies

$$\begin{bmatrix} 4 \\ -3 \end{bmatrix} \quad = \quad 2 \odot \begin{bmatrix} 3 \\ 0 \end{bmatrix} - 3 \odot \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

In other words the coordinates of vector $\vec{u}$ in basis B, equivalently its representation in basis B is

$$\mathcal{R}_B\left(\begin{bmatrix} 4 \\ -3 \end{bmatrix}\right) \quad = \quad \begin{pmatrix} 2 \\ -3 \end{pmatrix}_B .$$

Consider a different basis $D = \left\{\vec{d}_1, \vec{d}_2\right\}$ where

$$\vec{d}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \vec{d}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

the vector equation is

$$\begin{aligned}
\begin{bmatrix} 4 \\ -3 \end{bmatrix} &= x_1 \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \oplus x_2 \odot \begin{bmatrix} 3 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} -2\,x_1 + 2 \\ x_1 \end{bmatrix} \oplus \begin{bmatrix} x_2 + 2 \\ -x_2 \end{bmatrix} \\
&= \begin{bmatrix} -2\,x_1 + x_2 + 2 \\ x_1 - x_2 \end{bmatrix}
\end{aligned}$$

and its corresponding system of linear equation

$$\left(\begin{array}{rr|r} -2 & 1 & 2 \\ 1 & -1 & -3 \end{array}\right)$$

implies

$$\begin{bmatrix} 4 \\ -3 \end{bmatrix} = 1 \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \oplus 4 \odot \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

The coordinates of vector $\vec{u}$ in basis D, equivalently its representation in basis D is

$$\mathcal{R}_D\left(\begin{bmatrix} 4 \\ -3 \end{bmatrix}\right) = \begin{pmatrix} 1 \\ 4 \end{pmatrix}_D$$

Recall: In the vectors space $\mathbb{R}^3$ the subset

$$\mathbb{S} = \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mid 3X + 6Y = 2Z \right\}$$

is a subspace and

$$\mathbb{S} = \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} \right\rangle$$

For the vector $\vec{u} = \begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix}$ the vector equation

$$\begin{pmatrix} 4 \\ -3 \\ -3 \end{pmatrix} = x_1 \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$$

is equivalent to a system of linear equations with augmented matrix

$$\left(\begin{array}{rrr|r} 2 & 0 & 2 & 4 \\ -1 & 1 & 0 & -3 \\ 0 & 3 & 3 & -3 \end{array}\right)$$

that has infinitely many solutions. For example

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

In other words vector $\vec{u}$ cannot be represented in a unique way. The ambiguity is the reason to require linearly independent vectors in a basis.

Theorem 31. *In a vector space a subset* $\mathbf{B}$ *is a basis if and only if any vector can be represented in a unique way as a linear combination of the vectors in* $\mathbf{B}$.

Proof. If $\mathbf{B}$ is a basis then by definition every vector in the vector space is represented in a unique way. Conversely, if every vector in the vector space can be represented in a unique way as a linear combination of the vectors in $\mathbf{B}$, then by definition $\mathbf{B}$ spans the vector space. The zero vector can be represented as a linear combination of the vectors in $\mathbf{B}$ by taking all coefficients as zero and apply Theorem 8. Since by assumption there is only one way to represent every vector, this is the only way to represent the zero vector, which means the vectors in $\mathbf{B}$ are linearly independent and thus $\mathbf{B}$ is a basis. $\square$

2.10.1 Change of basis

Let $\vec{u} \in \mathbf{V}$ and let $\mathbf{B}$ and $\mathbf{E}$ be two bases for $\mathbf{V}$. Given the coordinates of $\vec{u}$ with respect to $\mathbf{E}$ what are its coordinates with respect to $\mathbf{B}$? We will illustrate the answer with a few examples.

Example: Recall from §2.10 bases $\mathbf{E}$ and $\mathbf{B}$ for $\mathbb{R}^3$.

$$\begin{aligned}
\mathbb{R}^3 = \langle \mathbf{E} \rangle &= \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \\
\langle \mathbf{B} \rangle &= \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle
\end{aligned}$$

For each basis vector $\mathbf{E}$ compute its representation in basis $\mathbf{B}$. For $\vec{e}_1$ solve

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + x_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

to obtain

$$\mathcal{R}_B(\vec{e}_1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

For $\vec{e}_2$ solve

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

to obtain

$$\mathcal{R}_B(\vec{e}_2) = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

For $\vec{e}_3$ solve

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

to obtain

$$\mathcal{R}_B(\vec{e}_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Write the representations in that order to a matrix

$$\mathcal{R}_{E\to B}(id) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

The result is the change of basis from basis **E** to basis **B**. For that matrix the following matrix equation is satisfied

$$\mathcal{R}_B(\vec{u}) = \mathcal{R}_{E\to B}(id)\,\mathcal{R}_E(\vec{u})$$

with the values from the example

$$\underbrace{\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}}_{\mathcal{R}_B(\vec{u})} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}}_{\mathcal{R}_{E\to B}(id)} \underbrace{\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}}_{\mathcal{R}_E(\vec{u})}.$$

Example: Recall from §2.10 bases **B** and **D** for the vectors space

$$\mathbf{S} = \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mid 3X + 6Y = 2Z \right\}.$$

For each basis vector $\mathbf{B}$ compute its representation in basis $\mathbf{D}$

$$\begin{aligned}
\mathbb{S} = \langle \mathbf{B} \rangle &= \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \right\rangle \\
\langle \mathbf{D} \rangle &= \left\langle \begin{pmatrix} 4 \\ -5 \\ -9 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix} \right\rangle
\end{aligned}$$

For $\vec{b}_1$ solve

$$\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 4 \\ -5 \\ -9 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix}$$

to obtain

$$\mathcal{R}_D\left(\vec{b}_1\right) = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

For $\vec{b}_2$ solve

$$\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} = x_1 \begin{pmatrix} 4 \\ -5 \\ -9 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix}$$

to obtain

$$\mathcal{R}_D\left(\vec{b}_2\right) = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Write the representations in that order to a matrix

$$\mathcal{R}_{B\to D}\,(id) = \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix}$$

The result is the change of basis from basis $\mathbf{B}$ to basis $\mathbf{D}$. For that matrix the following matrix equation is satisfied

$$\mathcal{R}_D(\vec{u}) = \mathcal{R}_{B\to D}\,(id)\,\mathcal{R}_B(\vec{u})$$

with the values from the example

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

Example: Recall from §2.10 bases $\mathbf{B}$ and $\mathbf{D}$ for the vectors space $\mathbf{P}_2$. For each basis vector $\mathbf{B}$ compute its representation in basis $\mathbf{D}$

$$\begin{aligned}
\mathbf{P}_2 = \langle \mathbf{B} \rangle &= \left\langle 6\,x^2 - 39\,x + 23, \qquad -5\,x^2 + 33\,x - 19, \qquad -x^2 + 7\,x - 4 \right\rangle \\
\langle \mathbf{D} \rangle &= \left\langle x^2 - 8\,x + 5, \qquad x^2 - 5\,x + 3, \qquad -x^2 + 7\,x - 4 \right\rangle
\end{aligned}$$

For $\vec{b}_1$ solve

$$6\,x^2 - 39\,x + 23 \quad = \quad \alpha_1\left(x^2 - 8\,x + 5\right) + \alpha_2\left(x^2 - 5\,x + 3\right) + \alpha_3\left(-x^2 + 7\,x - 4\right)$$

to obtain

$$\mathcal{R}_D\left(\vec{b}_1\right) = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$$

For $\vec{b}_2$ solve

$$-5\,x^2 + 33\,x - 19 \quad = \quad \alpha_1\left(x^2 - 8\,x + 5\right) + \alpha_2\left(x^2 - 5\,x + 3\right) + \alpha_3\left(-x^2 + 7\,x - 4\right)$$

to obtain

$$\mathcal{R}_D\left(\vec{b}_2\right) = \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix}$$

For $\vec{b}_3$ solve

$$-x^2 + 7\,x - 4 \quad = \quad \alpha_1\left(x^2 - 8\,x + 5\right) + \alpha_2\left(x^2 - 5\,x + 3\right) + \alpha_3\left(-x^2 + 7\,x - 4\right)$$

to obtain

$$\mathcal{R}_D\left(\vec{b}_3\right) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Write the representations in that order to a matrix

$$\mathcal{R}_{B\to D}\,(id) = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -3 & 4 & 1 \end{pmatrix}$$

The result is the change of basis from basis **B** to basis **D**. For that matrix the following matrix equation is satisfied

$$\mathcal{R}_D(\vec{u}) = \mathcal{R}_{B\to D}\,(id)\,\mathcal{R}_B(\vec{u})$$

with the values from the example

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

Example: Recall from §2.10 bases **B** and **D** the vector space **CVS** described in §2.2.

$$\begin{aligned} \mathbf{CVS} = \langle \mathbf{B} \rangle \quad &= \quad \left\langle \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\rangle \\ \langle \mathbf{D} \rangle \quad &= \quad \left\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\rangle \end{aligned}$$

For each basis vector $\mathbf{B}$ compute its representation in basis $\mathbf{D}$. For $\vec{b}_1$ solve

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \alpha_1 \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \oplus \alpha_2 \odot \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

to obtain

$$\mathcal{R}_D\left(\vec{b}_1\right) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

For $\vec{b}_2$ solve

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \alpha_1 \odot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \oplus \alpha_2 \odot \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

to obtain

$$\mathcal{R}_D\left(\vec{b}_2\right) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

Write the representations in that order to a matrix

$$\mathcal{R}_{B\to D}\left(id\right) = \begin{pmatrix} -1 & -1 \\ -1 & -2 \end{pmatrix}$$

The result is the change of basis from basis $\mathbf{B}$ to basis $\mathbf{D}$. For that matrix the following matrix equation is satisfied

$$\mathcal{R}_D(\vec{u}) = \mathcal{R}_{B\to D}\left(id\right)\mathcal{R}_B(\vec{u})$$

with the values from the example

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

2.11 Rank of a matrix

Theorem 32. *Let A be a square matrix for which there is a square matrix B such that $AB = I$. Then the columns of A are linearly independent.*

Proof. Let A be $k \times k$ matrix. Since A is invertible then there exists A^{-1} such that $AA^{-1} = I$. In the last multiplication it follows that the columns of I are linear combinations of the columns of A. Denote the columns of A as $\vec{a_1}, \ldots, \vec{a_k}$ and suppose by contradiction that $\vec{a_1}, \ldots, \vec{a_k}$ are linearly dependent. Let $\vec{b_1}, \ldots, \vec{b_m}$, where $m < k$ be a set with largest cardinality such that $\vec{b_1}, \ldots, \vec{b_m}$ are linearly independent. Then the columns of I are linear combinations of $\vec{b_1}, \ldots, \vec{b_m}$. This follows from the fact that $\vec{a_1}, \ldots, \vec{a_k}$ are linear combinations of $\vec{b_1}, \ldots, \vec{b_m}$. Then by Theorem 20 the columns of I are linearly dependent, which is a contradiction with Theorem 15. $\square$

Recall Theorem 4

> If for a matrix A there exists matrices B and C such that $AB = I$ and $AC = 0$ then $C = 0$, where 0 is the zero matrix.

Here is an alternative proof using linear dependence and independence:

Proof. Let A be $k \times k$ matrix and assume by contradiction $AB = 0$ and $B \neq 0$. Let the j^{th} row of B contain a non-zero element. Denote the j^{th} row of B by $\vec{b}_j^{\,t}$. Then $A\vec{b}_j^{\,t} = \vec{0}$, which means that the columns of A are linear dependent contradicting Theorem 32. ☐

Theorem 33. *Let A be $n \times m$ matrix. Then the number of linearly independent rows equals the number of linearly independent columns.*

Proof. Suppose A's columns are spanned by $\vec{b}, \dots, \vec{b}_r$ then there is an $r \times m$ matrix C such that

$$[\vec{a}_1 \dots \vec{a}_m] = A = BC = [\vec{b}_1 \dots \vec{b}_r]C.$$

By properties of matrix multiplication the rows of A are linear combinations of the rows of C and therefore the rows of A contain at most r linearly independent rows. Thus the number of linear independent rows of A do not exceed the number of linear independent columns of A. Applying the same argument for the transpose of A we obtain that the number of linearly independent rows of A equals the number of linearly independent columns of A. ☐

Example: Illustration of the above proof for the matrix

$$A = \begin{pmatrix} 1 & -2 & 1 & 4 & -2 \\ 2 & -4 & 3 & 8 & -3 \\ 1 & -2 & 2 & 5 & -1 \\ 1 & -2 & 1 & 0 & -2 \end{pmatrix}$$

we have that

$$\overbrace{\begin{pmatrix} 1 & -2 & 1 & 4 & -2 \\ 2 & -4 & 3 & 8 & -3 \\ 1 & -2 & 2 & 5 & -1 \\ 1 & -2 & 1 & 0 & -2 \end{pmatrix}}^{A} = \overbrace{\begin{pmatrix} \underbrace{\begin{matrix} 1 \\ 2 \\ 1 \\ 1 \end{matrix}}_{\vec{b}_1} & \underbrace{\begin{matrix} 1 \\ 3 \\ 2 \\ 1 \end{matrix}}_{\vec{b}_2} & \underbrace{\begin{matrix} 4 \\ 8 \\ 5 \\ 0 \end{matrix}}_{\vec{b}_3} \end{pmatrix}}^{B} \overbrace{\begin{pmatrix} 1 & -2 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}}^{C}$$

In this case $n = 4$, $m = 5$ and $r = 3$.

Definition 42 (rank of a matrix)**.** *The rank of a matrix A is the number of linear independent columns of A denoted by* $\operatorname{rank}(A)$*.*

Example: For the matrix A from the previous example we have that the following linear dependence relations

$$\begin{pmatrix} -2 \\ -4 \\ -2 \\ -2 \end{pmatrix} = -2\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -2 \\ -3 \\ -1 \\ -2 \end{pmatrix} = -3\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix}$$

and the only solution to

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix} x_2 + \begin{pmatrix} 4 \\ 8 \\ 5 \\ 0 \end{pmatrix} x_3$$

is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which means that the matrix A has three linearly independent columns. In terms of rows we have the following relation

$$\begin{aligned} \overbrace{(1,\ -2,\ 2,\ 5,\ -1)}^{\text{row 3}} &= \frac{-3}{4}\overbrace{(1,\ -2,\ 1,\ 4,\ -2)}^{\text{row 1}} \\ &\quad + \overbrace{(2,\ -4,\ 3,\ 8,\ -3)}^{\text{row 2}} \\ &\quad + \frac{-1}{4}\overbrace{(1,\ -2,\ 1,\ 0,\ -2)}^{\text{row 4}} \end{aligned}$$

and the only solution to

$$\begin{aligned} (0,\ 0,\ 0,\ 0,\ 0) &= (1,\ -2,\ 1,\ 4,\ -2)\, x_1 \\ &\quad + (2,\ -4,\ 3,\ 8,\ -3)\, x_2 \\ &\quad + (1,\ -2,\ 1,\ 0,\ -2)\, x_3 \end{aligned}$$

is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which means the matrix A has three linearly independent rows.

Theorem 34. *Suppose the rows (columns) of square A are linearly independent then A can be written as a product of elementary matrices.*

Proof. The columns of A form a basis for the n-dimensional vector space $\mathbb{K}^n$. Indeed if $\vec{e}_i$ is not in the span of the columns of $A = (\vec{c}_1, \dots, \vec{c}_n)$ then $\{\vec{c}_1, \dots, \vec{c}_n, \vec{e}_i\}$ are linearly independent (why?), which means in $\mathbb{K}^n$ we have found $n+1$ linearly independent vectors contradiction with the fact that the dimension of $\mathbb{K}^n$ is a n. Therefore the standard basis can be represented as linear combinations of the columns of A. We can solve $A\vec{x} = \vec{e}_i$ so there exists $\vec{c}_i$ such that $A\vec{c}_i = \vec{e}_i$ and by setting $B = (\vec{c}_1, \dots, \vec{c}_n)$ we obtain that $AB = I$ thus A is invertible. Using the matrix representation of Gaussian operations we can write $A = B^{-1} = E_m \dots E_1$ where each E_i is an elementary matrix. □

Theorem 35. *A system of linear equation $Ax = b$ has a solution if and only if the rank of the matrix of the system A equals the rank of the augmented matrix $(A|b)$ or* $\operatorname{rank}(A) = \operatorname{rank}(A|b)$.

Proof. By Theorem 24 $Ax = b$ has a solution if and only if b is in the span of the columns of A. By Theorem 25 the vector b is in span of the columns of A if and only if the span of the columns of A and the span of the columns of $(A|b)$ are equal. Since the number of linear independent columns of a matrix is by definition the rank of the matrix the result follows. □

Theorem 36. *$Ax = b$ has a solution if and only if $A^T y = 0 \Rightarrow b^T y = 0$*

Proof. Let $\vec{a}_i$ denote the ith column of matrix A that is $A = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$.

[⇒] Theorem 24 then if $Ax = b$ has a solution then $\vec{b}$ is in the span of the columns of A and therefore $\vec{b} = \alpha_1 \vec{a}_1 + \cdots \alpha_n \vec{a}_n$ for some constants $\alpha_1, \dots, \alpha_n$. Let $\vec{y}$ be any vector such that $\vec{y}^T A = \vec{0}$ this means that $\vec{y}^T \vec{a}_j = 0$ for all j by properties of matrix multiplication. Then $\vec{y}^T \vec{b} = \vec{y}^T (\alpha_1 \vec{a}_1 + \cdots + \alpha_n \vec{a}_n) = \alpha_1 \vec{y}^T \vec{a}_1 + \cdots + \alpha_n \vec{y}^T \vec{a}_n) = \alpha_1 0 + \cdots + \alpha_n 0 = 0$

[⇐] Assume $A^T y = 0 \Rightarrow b^T y = 0$ then

$$\begin{pmatrix} A^T \\ b^T \end{pmatrix} y = 0 \Leftrightarrow \begin{pmatrix} A^T \end{pmatrix} y = 0$$

Let $B = \begin{pmatrix} A^T \\ b^T \end{pmatrix}$ In particular this means that that the number of linearly independent columns of B equals the number of linear independent columns of A^T. Indeed if the $A^T y = \vec{0}$ then by our assumption $By = \vec{0}$. Thus for any set of linear dependent columns of A^T the corresponding set of columns of B is also linearly dependent. So the number of linear independent columns of B is less than or equal to the number of linear independent columns of A^T. On the other hand if a set of columns in A^T are linearly independent then the corresponding set of columns of B is also linearly independent as adding an equation to a system of linear equation cannot increase the number of solutions (and for linear independence we have only one solution namely all zeroes). Thus the we have $\mathrm{rank}(A) = \mathrm{rank}(A^T) = \mathrm{rank}(B) = \mathrm{rank}(B^T) = \mathrm{rank}(B) = \mathrm{rank}(A|b)$ and by Theorem 35 the result follows. $\square$

Chapter 3

Linear Transformations

3.1 Basic definitions

Definition 43 (linear map)**.** *Let* $\mathbf{U}$ *and* $\mathbf{W}$ *be two vector spaces. A function* ϕ : $\mathbf{U} \to \mathbf{W}$ *is* linear map (homomorphism) *if*

1. $\forall \vec{u}, \vec{v} \in \mathbf{U}, \phi(\vec{u} + \vec{v}) = \phi(\vec{u}) + \phi(\vec{v})$
2. $\forall c \in \mathbb{K}, \forall \vec{u} \in \mathbf{U}, \phi(c\vec{u}) = c\phi(\vec{u})$

In this case $\mathbf{U}$ *is called the* domain *and* $\mathbf{W}$ *is called the* co-domain *of* ϕ*.*

Theorem 37. *ϕ is a linear map if and only if $\phi(\alpha\vec{u} + \beta\vec{v}) = \alpha\phi(\vec{u}) + \beta\phi(\vec{v})$.*

Proof. Assume

$$\forall \alpha, \beta \in \mathbb{K}, \forall \vec{u}, \vec{v} \in \mathbf{U}, \phi(\alpha\vec{u} + \beta\vec{v}) = \alpha\phi(\vec{u}) + \beta\phi(\vec{v}).$$

With $\alpha = 1$ and $\beta = 1$ for all $\vec{u}, \vec{v} \in \mathbf{U}$ we get

$$\begin{aligned} \phi(\vec{u} + \vec{v}) &= \phi(1\vec{u} + 1\vec{v}) \\ &= 1\phi(\vec{u}) + 1\phi(\vec{v}) \\ &= \phi(\vec{u}) + \phi(\vec{v}) \end{aligned}$$

Thus the first condition of Definition 43 is satisfied.

With $\beta = 0$ and $\vec{v} = \vec{u}$ for all $\vec{u}, \vec{v} \in \mathbf{U}$ we get

$$\begin{aligned} \phi(\alpha\vec{u}) &= \phi\left(\alpha\vec{u} + \vec{0}_U\right) \\ &= \phi(\alpha\vec{u} + 0\vec{0}_U) \\ &= \alpha\phi(\vec{u}) + 0\phi\left(\vec{0}_U\right) \\ &= \alpha\phi(\vec{u}) + \vec{0}_W \\ &= \alpha\phi(\vec{u}) \end{aligned}$$

Thus the second condition of Definition 43 is satisfied.

For the converse assume

1. $\forall \vec{u}, \vec{v} \in \mathbf{U}, \phi(\vec{u}+\vec{v}) = \phi(\vec{u}) + \phi(\vec{v})$
2. $\forall c \in \mathbb{K}, \forall \vec{u} \in \mathbf{U}, \phi(c\vec{u}) = c\phi(\vec{u})$

then

$$\begin{aligned}\phi(\alpha\vec{u}+\beta\vec{v}) &= \phi\left(\alpha\vec{u}\right) + \phi\left(\beta\vec{v}\right) \\ &= \alpha\phi(\vec{u}) + \beta\phi\left(\vec{v}\right)\end{aligned}$$

Which completes the argument. □

We will use the condition in Theorem 37 as definition for linear map in some of the arguments, without explicitly stating that it is equivalent to Definition 43.

Definition 44. *[isomorphism] Let* $\mathbf{U}$ *and* $\mathbf{W}$ *be two vector spaces. A function* $\phi : \mathbf{U} \to \mathbf{W}$ *is an* isomorphism *between* $\mathbf{U}$ *and* $\mathbf{W}$ *if*

1. *ϕ is one-to-one and onto (correspondence)*
   (a) *onto* $\forall \vec{w} \in \mathbf{W}, \exists \vec{v} \in \mathbf{U} : \phi(\vec{v}) = \vec{w}$
   (b) *1-1* $\forall \vec{u}, \vec{v} \in \mathbf{U}, \phi(\vec{u}) = \phi(\vec{v}) \Rightarrow \vec{u} = \vec{v}$
2. *ϕ is a linear map.*

We write $\mathbf{U} \cong \mathbf{W}$ *if there is an isomorphism between* $\mathbf{U}$ *and* $\mathbf{W}$.

Definition 45. *A linear map (homomorphism) from a vector space* $\mathbf{V}$ *to itself is called a* linear transformation*.*

Definition 46 (automorphism)**.** *An isomorphism from* $\mathbf{U}$ *to itself is called an automorphism.*

3.1.1 Note on terminology

Let $\phi : \mathbf{U} \to \mathbf{W}$ then

homomorphism: also linear (map/function) $\phi\left(\alpha\vec{u}+\beta\vec{v}\right) = \alpha\phi\left(\vec{u}\right) + \beta\phi\left(\vec{v}\right)$;

transformation: $\mathbf{U} = \mathbf{W}$ that is domain and co-domain are the same

isomorphism: ϕ is all of homomorphism, one-to-one and onto

automorphism: ϕ is both isomorphism and transformation

3.2 Examples

3.2.1 Reflection

Reflection along the xy-plane.

$$\phi\left(\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_0 \\ x_1 \\ -x_2 \end{pmatrix}$$

Onto: Given any $\begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$ the system of linear equations

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$$

always has a solution for example, so the map is onto.

One-to-one: Given any $\begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$ the system of linear equations

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = (y_0,\, y_1,\, y_2)$$

always has a *unique* solution:

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ -y_2 \end{pmatrix}$$

so the map is one-to-one.

Linearity: We have

$$\begin{aligned}
\phi\left(a\vec{u}+b\vec{v}\right) &= \phi\left(a\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} + b\begin{pmatrix} v_0 \\ v_1 \\ v_2 \end{pmatrix}\right) \\
&= \phi\left(\begin{pmatrix} au_0+bv_0 \\ au_1+bv_1 \\ au_2+bv_2 \end{pmatrix}\right) \\
&= \begin{pmatrix} au_0+bv_0 \\ au_1+bv_1 \\ -au_2-bv_2 \end{pmatrix} \\
&= a\begin{pmatrix} u_0 \\ u_1 \\ -u_2 \end{pmatrix} + b\begin{pmatrix} v_0 \\ v_1 \\ -v_2 \end{pmatrix} \\
&= a\,\phi\left(\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}\right) + b\,\phi\left(\begin{pmatrix} v_0 \\ v_1 \\ v_2 \end{pmatrix}\right) \\
&= a\,\phi\left(\vec{u}\right) + b\,\phi\left(\vec{v}\right)
\end{aligned}$$

in other words

$$\phi\left(a\vec{u}+b\vec{v}\right) = a\,\phi\left(\vec{u}\right) + b\,\phi\left(\vec{v}\right)$$

thus the function is a linear map.

This function is

1. one-to-one
2. onto
3. homomorphism
4. linear transformation
5. isomorphism
6. automorphism

With range diagonal matrices

If we have the map

$$\phi\left(\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_0 & 0 & 0 \\ 0 & x_1 & 0 \\ 0 & 0 & -x_2 \end{pmatrix}$$

from $\mathbb{R}^3$ to the set of 3×3 *diagonal* matrices, similar to the above the new function is

1. one-to-one
2. onto
3. homomorphism
4. isomorphism

but it is *not*

1. linear transformation
2. automorphism

With range 3×3 **matrices**

If we have the map

$$\phi\left(\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_0 & 0 & 0 \\ 0 & x_1 & 0 \\ 0 & 0 & -x_2 \end{pmatrix}$$

from $\mathbb{R}^3$ to the set of *all* 3×3 matrices, similar to the above the new function is

1. one-to-one
2. homomorphism

but it is *not*

1. onto
2. isomorphism
3. linear transformation
4. automorphism

3.2.2 Example: exponential coordinate

$$\phi\left(\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}\right) = \begin{pmatrix} e^{u_0} \\ u_1 \\ u_1 + u_2 \end{pmatrix}$$

Onto: Given $\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}$ from calculus $e^{u_0} = -1$ has no solution so there is no vector $\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}$ such that

$$\phi\left(\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

thus the map is not onto.

One-to-one: From calculus $e^{u_0} = y$ is one to one function. So if

$$\begin{pmatrix} e^{u_0} \\ u_1 \\ u_1 + u_2 \end{pmatrix} = (y_0,\, y_1,\, y_2)$$

has a solution, that solution is *unique*, so the map is one-to-one.

Linearity: We have

$$\begin{aligned}
\phi\left(a\vec{u} + b\vec{v}\right) &= \phi\left(2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}\right) \\
&= \phi\left(\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}\right) \\
&= \begin{pmatrix} e^2 \\ 0 \\ 0 \end{pmatrix} \\
&\neq \begin{pmatrix} 2\,e \\ 0 \\ 0 \end{pmatrix} \\
&= 2\begin{pmatrix} e \\ 0 \\ 0 \end{pmatrix} + 0\begin{pmatrix} e^2 \\ 0 \\ 0 \end{pmatrix} \\
&= 2\phi\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) + 0\phi\left(\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}\right)
\end{aligned}$$

in other words

$$\phi\left(a\vec{u} + b\vec{v}\right) \neq a\,\phi\left(\vec{u}\right) + b\,\phi\left(\vec{v}\right)$$

thus the function is not a linear map.

This function is

1. one-to-one
2. not onto
3. not homomorphism
4. not linear transformation
5. not isomorphism
6. not automorphism

3.2.3 Example: polynomial coordinate

$$\phi\left(\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}\right) = \begin{pmatrix} (u_0+1)(u_0-1)u_0 \\ u_2 \\ u_1 \end{pmatrix}$$

Onto: Given $\begin{pmatrix} v_0 \\ v_1 \\ v_2 \end{pmatrix}$ from calculus $x^3 - x = y$ is an onto function so for any y_0 there is u_0 such that $u_0^3 - u_0 = y_0$ and

$$\phi^{-1}\left(\begin{pmatrix} v_0 \\ v_1 \\ v_2 \end{pmatrix}\right) = \begin{pmatrix} u_0 \\ y_2 \\ y_1 \end{pmatrix}$$

thus the map is onto.

One-to-one: From calculus $x^3 - x = y$ is not one-to-one function. For example

$$\phi\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \phi\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) =$$

so the map is *not* one-to-one.

Linearity: We have

$$\begin{aligned}
\phi\left(a\vec{u}+b\vec{v}\right) &= \phi\left(2\begin{pmatrix}1\\0\\0\end{pmatrix}+0\begin{pmatrix}1\\0\\0\end{pmatrix}\right)\\
&= \phi\left(\begin{pmatrix}2\\0\\0\end{pmatrix}\right)\\
&= \begin{pmatrix}6\\0\\0\end{pmatrix}\\
&\neq \begin{pmatrix}0\\0\\0\end{pmatrix}\\
&= 2\begin{pmatrix}0\\0\\0\end{pmatrix}+0\begin{pmatrix}0\\0\\0\end{pmatrix}\\
&= 2\phi\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right)+0\phi\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right)
\end{aligned}$$

in other words

$$\phi\left(a\vec{u}+b\vec{v}\right)\neq a\,\phi\left(\vec{u}\right)+b\,\phi\left(\vec{v}\right)$$

thus the function is not a linear map.

This function is

1. not one-to-one
2. onto
3. not homomorphism
4. not linear transformation
5. not isomorphism
6. not automorphism

3.2.4 Polynomials to upper triangular matrices:

Consider the map from polynomials of degree two to upper triangular 2×2 matrices.

$$\phi\left(u_2x^2+u_1x+u_0\right)=\begin{pmatrix} u_0+u_1 & -u_1+u_2\\ 0 & u_0+2\,u_2\end{pmatrix}$$

Let S denote the following system of linear equations in matrix form

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$$

onto For any vector $\begin{pmatrix} y_0 & y_1 \\ 0 & y_2 \end{pmatrix}$ the system of linear equations S always has a solution for any values $\begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$ so the map is onto.

0ne-to-one For any vector $\begin{pmatrix} y_0 & y_1 \\ 0 & y_2 \end{pmatrix}$ the system of linear equations S always has a *unique* solution for any values $\begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$ so the map is one-to-one.

Linearity: We have

$$\begin{aligned}
\phi\left(a\vec{u} + b\vec{v}\right) &= \phi\left(a\left(u_2x^2 + u_1x + u_0\right) + b\left(v_2x^2 + v_1x + v_0\right)\right) \\
&= \phi\left((au_2 + bv_2)x^2 + au_0 + bv_0 + (au_1 + bv_1)x\right) \\
&= \begin{pmatrix} au_0 + au_1 + bv_0 + bv_1 & -au_1 + au_2 - bv_1 + bv_2 \\ 0 & au_0 + 2\,au_2 + bv_0 + 2\,bv_2 \end{pmatrix} \\
&= \begin{pmatrix} a(u_0 + u_1) + b(v_0 + v_1) & -a(u_1 - u_2) - b(v_1 - v_2) \\ 0 & a(u_0 + 2\,u_2) + b(v_0 + 2\,v_2) \end{pmatrix} \\
&= a\begin{pmatrix} u_0 + u_1 & -u_1 + u_2 \\ 0 & u_0 + 2\,u_2 \end{pmatrix} + b\begin{pmatrix} v_0 + v_1 & -v_1 + v_2 \\ 0 & v_0 + 2\,v_2 \end{pmatrix} \\
&= a\phi\left(u_2x^2 + u_1x + u_0\right) + b\phi\left(v_2x^2 + v_1x + v_0\right)
\end{aligned}$$

in other words

$$\phi\left(a\vec{u} + b\vec{v}\right) = a\,\phi\left(\vec{u}\right) + b\,\phi\left(\vec{v}\right)$$

thus the function is a linear map.

This function is

1. one-to-one
2. onto
3. homomorphism (linear map)
4. not linear transformation
5. isomorphism

6. not automorphism

If in the above example we change the co-domain to the set of *all* square matrices instead of upper triangular square matrices, the function will be

1. one-to-one
2. not onto
3. homomorphism (linear map)
4. not linear transformation
5. not isomorphism
6. not automorphism

3.2.5 Example: $\mathcal{M}_{2\times 3}$ to $\mathcal{M}_{3\times 2}$ matrices

Domain 2×3 matrices, co-domain 3×2 matrices

$$\begin{aligned}
\phi(\vec{u}) &= \phi\left(\begin{pmatrix} u_0 & u_1 & u_2 \\ u_3 & u_4 & u_5 \end{pmatrix}\right) \\
&= \vec{w} = \begin{pmatrix} y_0 & y_1 \\ y_2 & y_3 \\ y_4 & y_5 \end{pmatrix} \\
&= \begin{pmatrix} u_0 + u_2 + u_5 & u_1 - 2\,u_3 \\ u_0 - 3\,u_1 + 2\,u_3 & 2\,u_1 + u_5 \\ u_1 + u_2 - 2\,u_3 & 2\,u_1 - 4\,u_3 \end{pmatrix}
\end{aligned}$$

Let S denote the following system of linear equations in matrix form

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 & 0 & 0 \\ 1 & -3 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -2 & 0 & 0 \\ 0 & 2 & 0 & -4 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}$$

onto For vector $\begin{pmatrix} 3 & 0 \\ -2 & 0 \\ 0 & 1 \end{pmatrix}$ the corresponding system of linear equations S

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 & 0 & 0 \\ 1 & -3 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -2 & 0 & 0 \\ 0 & 2 & 0 & -4 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

has no solution, so the map is not onto.

0ne-to-one For vectors

$$\vec{u} = \begin{pmatrix} 4 & 2 & 1 \\ 1 & 0 & -4 \end{pmatrix} \qquad \vec{v} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -5 & 0 \end{pmatrix}$$

we have

$$\phi\left(\begin{pmatrix} 4 & 2 & 1 \\ 1 & 0 & -4 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} = \phi\left(\begin{pmatrix} 0 & 0 & 1 \\ 0 & -5 & 0 \end{pmatrix}\right)$$

so the map is not one-to-one.

Linearity: We have

$$\begin{aligned}
\phi\left(a\vec{u} + b\vec{v}\right) &= \phi\left(a\begin{pmatrix} u_0 & u_1 & u_2 \\ u_3 & u_4 & u_5 \end{pmatrix} + b\begin{pmatrix} v_0 & v_1 & v_2 \\ v_3 & v_4 & v_5 \end{pmatrix}\right) \\
&= \phi\left(\begin{pmatrix} au_0 + bv_0 & au_1 + bv_1 & au_2 + bv_2 \\ au_3 + bv_3 & au_4 + bv_4 & au_5 + bv_5 \end{pmatrix}\right) \\
&= \begin{pmatrix} au_0 + au_2 + au_5 + bv_0 + bv_2 + bv_5 & au_1 - 2\,au_3 + bv_1 - 2\,bv_3 \\ au_0 - 3\,au_1 + 2\,au_3 + bv_0 - 3\,bv_1 + 2\,bv_3 & 2\,au_1 + au_5 + 2\,bv_1 + bv_5 \\ au_1 + au_2 - 2\,au_3 + bv_1 + bv_2 - 2\,bv_3 & 2\,au_1 - 4\,au_3 + 2\,bv_1 - 4\,bv_3 \end{pmatrix} \\
&= \begin{pmatrix} a(u_0 + u_2 + u_5) + b(v_0 + v_2 + v_5) & a(u_1 - 2\,u_3) + b(v_1 - 2\,v_3) \\ a(u_0 - 3\,u_1 + 2\,u_3) + b(v_0 - 3\,v_1 + 2\,v_3) & a(2\,u_1 + u_5) + b(2\,v_1 + v_5) \\ a(u_1 + u_2 - 2\,u_3) + b(v_1 + v_2 - 2\,v_3) & 2\,a(u_1 - 2\,u_3) + 2\,b(v_1 - 2\,v_3) \end{pmatrix} \\
&= a\begin{pmatrix} u_0 + u_2 + u_5 & u_1 - 2\,u_3 \\ u_0 - 3\,u_1 + 2\,u_3 & 2\,u_1 + u_5 \\ u_1 + u_2 - 2\,u_3 & 2\,u_1 - 4\,u_3 \end{pmatrix} + b\begin{pmatrix} v_0 + v_2 + v_5 & v_1 - 2\,v_3 \\ v_0 - 3\,v_1 + 2\,v_3 & 2\,v_1 + v_5 \\ v_1 + v_2 - 2\,v_3 & 2\,v_1 - 4\,v_3 \end{pmatrix} \\
&= a\phi\left(\begin{pmatrix} u_0 & u_1 & u_2 \\ u_3 & u_4 & u_5 \end{pmatrix}\right) + b\phi\left(\begin{pmatrix} v_0 & v_1 & v_2 \\ v_3 & v_4 & v_5 \end{pmatrix}\right)
\end{aligned}$$

in other words

$$\phi\left(a\vec{u} + b\vec{v}\right) = a\,\phi\left(\vec{u}\right) + b\,\phi\left(\vec{v}\right)$$

thus the function is a linear map.

The above example is

1. not one-to-one
2. not onto
3. homomorphism (linear map)
4. not linear transformation
5. not isomorphism
6. not automorphism

With range 2×3 matrices

If the above example is modified to

$$\begin{aligned}
\phi(\vec{u}) &= \phi\left(\begin{pmatrix} u_0 & u_1 & u_2 \\ u_3 & u_4 & u_5 \end{pmatrix}\right) \\
&= \vec{w} = \begin{pmatrix} y_0 & y_2 & y_4 \\ y_1 & y_3 & y_5 \end{pmatrix} \\
&= \begin{pmatrix} u_0 + u_2 + u_5 & u_0 - 3\,u_1 + 2\,u_3 & u_1 + u_2 - 2\,u_3 \\ u_1 - 2\,u_3 & 2\,u_1 + u_5 & 2\,u_1 - 4\,u_3 \end{pmatrix}
\end{aligned}$$

the resulting map is

1. not one-to-one
2. not onto
3. homomorphism (linear map)
4. linear transformation
5. not isomorphism
6. not automorphism

3.3 Isomorphic Vector Spaces

Theorem 38. *The representation map from a vector space* $\mathbf{V}$ *with basis* $\mathbf{B}$ *is a vector space with basis* $\vec{b}_1, \ldots, \vec{b}_d$ *to the vector space of standard column vectors with* d *components* $\mathbb{K}^d$ *is an isomorphism.*

$$\mathcal{R}_B : \mathbf{V} \to \mathbb{K}^d \qquad \mathcal{R}_B(\vec{u}) = \mathcal{R}_B(\alpha_1\vec{b}_1 + \cdots + \alpha_d\vec{b}_d) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix}_B$$

Proof. **onto** Let $\begin{pmatrix} c_1 \\ \vdots \\ c_d \end{pmatrix} \in \mathbb{K}^d$ consider $\vec{v} = c_1\vec{b}_1 + \cdots + c_d\vec{b}_d$. We have $\mathcal{R}_B(\vec{v}) = \begin{pmatrix} c_1 \\ \vdots \\ c_d \end{pmatrix}$

1-1 Suppose $\mathcal{R}_B(\vec{u}) = \begin{pmatrix} c_1 \\ \vdots \\ c_d \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix} = \mathcal{R}_B(\vec{v})$. Then $\vec{u} = c_1\vec{b}_1 + \cdots + c_d\vec{b}_d = a_1\vec{b}_1 + \cdots + a_d\vec{b}_d = \vec{v}$.

linear 1.

$$\begin{aligned} \mathcal{R}_B(\alpha\vec{u}) &= \mathcal{R}_B(\alpha c_1\vec{b}_1 + \cdots + \alpha c_d\vec{b}_d) \\ &= \begin{pmatrix} \alpha c_1 \\ \vdots \\ \alpha c_d \end{pmatrix} = \alpha \begin{pmatrix} c_1 \\ \vdots \\ c_d \end{pmatrix} \\ &= \alpha\mathcal{R}_B(\vec{u}) \end{aligned}$$

2.

$$\begin{aligned} \mathcal{R}_B(\vec{u}+\vec{v}) &= \mathcal{R}_B\left((c_1+a_1)\vec{b}_1 + \cdots + (c_d+a_d)\vec{b}_d\right) \\ &= \begin{pmatrix} c_1+a_1 \\ \vdots \\ c_d+a_d \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_d \end{pmatrix} + \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix} \\ &= \mathcal{R}_B(\vec{u}) + \mathcal{R}_B(\vec{v}) \end{aligned}$$

□

Theorem 39. *If $\phi : \mathbf{V} \to \mathbf{W}$ is an isomorphism then $\phi^{-1} : \mathbf{W} \to \mathbf{V}$ is also an isomorphism.*

Proof. An isomorphism is a correspondence between the sets so ϕ has an inverse function $\phi^{-1}\mathbf{W} \to \mathbf{V}$ which is also 1-1 and onto Since ϕ preserves linear combinations, so also does ϕ^{-1}. Let $\vec{w}_1, \vec{w}_2 \in \mathbf{W}$. Since ϕ is onto there are $\vec{v}_1, \vec{v}_2 \in \mathbf{V}$ such that $\vec{w}_1 = \phi(\vec{v}_1)$ and $\vec{w}_2 = \phi(\vec{v}_2)$. Then

$$\begin{aligned} f^{-1}(c_1 \cdot \vec{w}_1 + c_2 \cdot \vec{w}_2) &= f^{-1}\big(c_1 \cdot f(\vec{v}_1) + c_2 \cdot f(\vec{v}_2)\big) \\ &= f^{-1}\big(f\big(c_1\vec{v}_1 + c_2\vec{v}_2)\big) \\ &= c_1\vec{v}_1 + c_2\vec{v}_2 \\ &= c_1 \cdot f^{-1}(\vec{w}_1) + c_2 \cdot f^{-1}(\vec{w}_2) \end{aligned}$$

since $f^{-1}(\vec{w}_1) = \vec{v}_1$ and $f^{-1}(\vec{w}_2) = \vec{v}_2$. □

Theorem 40. *Isomorphism is an equivalence relation*

Proof. **reflexive** identity map

symmetric by Theorem 39

transitive standard argument from calculus

□

Theorem 41. *If two vector spaces are isomorphic then they have the same dimension.*

Proof. ☐

Theorem 42. *If two vector spaces (over the same field) have the same dimension then they are isomorphic.*

Proof. by the representation map both are isomorphic to $\mathbb{K}^d$, by transitivity isomorphic to each other. ☐

3.4 Properties of linear maps

Theorem 43. *For a linear map $\phi : \mathbf{V} \to \mathbf{W}$ we have that $\phi(\vec{0}_V) = \vec{0}_W$.*

Proof. Self study. ☐

Theorem 44. *If A is a linear transformation then it maps linearly dependent vectors to linearly dependent vectors*

Proof. Self study. ☐

Theorem 45. *If A is a linear transformation then it maps spanning sets to spanning sets.*

Proof. Self study. ☐

3.5 Linear extensions

In a standard calculus course you learn that in the Euclidean plane given any two points you can define a line. That is given two pairs of numbers you can create a unique equation $y = Ax + B$ that describes a line. The same idea can be generalized to linear maps. That is you can describe a linear map using just the definition of the map on any basis. Comparing with the line example you need the x-coordinates of the points are your basis and the y coordinates help identify the coefficients A and B. For linear maps the x-coordinates are the basis vectors, the y coordinates are the images of those x coordinates, just as if you plug in the x value of a point P to the line equation you get the corresponding y coordinates.

Theorem 46. *A homomorphism is determined by its action on a basis: if $\mathbf{V}$ is a vector space with basis $\vec{b}_1, \ldots, \vec{b}_n$ and $\mathbf{W}$ is a vector space with elements $\vec{w}_1, \ldots, \vec{w}_n$ (perhaps not distinct elements) then there exists a homomorphism from $\phi : \mathbf{V} \to \mathbf{W}$ such that $\phi(\vec{b}_i) = \vec{w}_i$, and that homomorphism is unique.*

Proof. **well-defined** let $\vec{v} \in \mathbf{V}$ and let $\vec{v} = v_1\vec{b}_1 + \cdots + v_n\vec{b}_n$. Define the associated output by using the same coordinates $\phi(\vec{v}) = v_1\vec{w}_1 + \cdots + v_n\vec{w}_n$. This is well defined because, with respect to the basis, since the representation of each vector $\vec{v}$ in the domain is unique for any given basis.

homomorphism This map is a homomorphism because it preserves linear combinations: where $\vec{u} = u_1\vec{b}_1 + \cdots + u_n\vec{b}_n$ and $\vec{v} = v_1\vec{b}_1 + \cdots + v_n\vec{b}_n$, here is the calculation.

$$\begin{aligned}
\phi\left(s\vec{u} + t\vec{v}\right) &= \phi\left((su_1 + tv_1)\vec{b}_1 + \cdots + (su_n + tv_n)\vec{b}_n \right) \\
&= (su_1 + tv_1)\phi\left(\vec{b}_1\right) + \cdots + (su_n + tv_n)\phi\left(\vec{b}_n\right) \\
&= (su_1 + tv_1)\vec{w}_1 + \cdots + (su_n + tv_n)\vec{w}_n \\
&= s\phi\left(\vec{u}\right) + \phi\left(\vec{v}\right)
\end{aligned}$$

unique This map is unique because if $\hat{\phi} : V \to W$ is another homomorphism such that $\hat{\phi}(\vec{b}_i) = \vec{w}_i$ for each i then ϕ and $\hat{\phi}$ agree on all of the vectors in the domain.

$$\begin{aligned}
\hat{\phi}(\vec{v}) &= \hat{\phi}(c_1\vec{b}_1 + \cdots + c_n\vec{b}_n) \\
&= c_1\hat{\phi}(\vec{b}_1) + \cdots + c_n\hat{\phi}(\vec{b}_n) \\
\\
&= c_1\vec{w}_1 + \cdots + c_n\vec{w}_n \\
&= \phi(\vec{v})
\end{aligned}$$

They have the same action so they are the same function.

□

Definition 47. *Let* $\mathbf{V}$ *and* $\mathbf{W}$ *be two vector spaces and let* $\mathbf{B} = \{\vec{b}_1, \ldots, \vec{b}_d\}$ *be a basis for* $\mathbf{V}$. *A function* f *defined on the basis* $\mathbf{B}$ *with* $f : \mathbf{B} \to \mathbf{W}$ *is* extended linearly *to a function* $\phi : \mathbf{V} \to \mathbf{W}$ *if* $\forall \vec{u} \in \mathbf{V}$ *with* $\vec{u} = u_1\vec{b}_1 + \cdots + u_d\vec{b}_d$, *the action of* ϕ *is defined as*

$$\phi(\vec{u}) = \phi(u_1\vec{b}_1 + \cdots + u_d\vec{b}_d) = u_1 f(\vec{b}_1) + \cdots + u_d f(\vec{b}_d)$$

The function ϕ is naturally a homomorphism.

3.5.1 $\mathbb{R}^2 \to \mathbb{R}^4$ example

Suppose you are given:

$$f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 2 \end{pmatrix} \qquad f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

then

$$\begin{aligned}
\phi\left(\left(\begin{array}{c} x_0 \\ x_1 \end{array}\right)\right) &= \phi\left(x_0\left(\begin{array}{c} 1 \\ 0 \end{array}\right) + x_1\left(\begin{array}{c} 0 \\ 1 \end{array}\right)\right) \\
&= x_0 f\left(\left(\begin{array}{c} 1 \\ 0 \end{array}\right)\right) + x_1 f\left(\left(\begin{array}{c} 0 \\ 1 \end{array}\right)\right) \\
&= x_0\left(\begin{array}{c} 2 \\ 0 \\ 1 \\ 2 \end{array}\right) + x_1\left(\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array}\right) = \left(\begin{array}{r} 2x_0 + x_1 \\ x_1 \\ x_0 + x_1 \\ 2x_0 + x_1 \end{array}\right)
\end{aligned}$$

for the vector $\left(\begin{array}{c} x_0 \\ x_1 \end{array}\right) = \left(\begin{array}{c} 2 \\ 3 \end{array}\right)$ we compute

$$\begin{aligned}
\phi\left(\left(\begin{array}{c} 2 \\ 3 \end{array}\right)\right) &= \phi\left(2\left(\begin{array}{c} 1 \\ 0 \end{array}\right) + 3\left(\begin{array}{c} 0 \\ 1 \end{array}\right)\right) \\
&= 2f\left(\left(\begin{array}{c} 1 \\ 0 \end{array}\right)\right) + 3f\left(\left(\begin{array}{c} 0 \\ 1 \end{array}\right)\right) \\
&= 2\left(\begin{array}{c} 2 \\ 0 \\ 1 \\ 2 \end{array}\right) + 3\left(\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array}\right) = \left(\begin{array}{c} 7 \\ 3 \\ 5 \\ 7 \end{array}\right)
\end{aligned}$$

Note: in general we can extend any set linearly as long as within the domain of the function f linearity is preserved:

$$f\left(\left(\begin{array}{c} 1 \\ 0 \end{array}\right)\right) = \left(\begin{array}{c} 2 \\ 0 \\ 1 \\ 2 \end{array}\right) \qquad f\left(\left(\begin{array}{c} 0 \\ 1 \end{array}\right)\right) = \left(\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array}\right) \qquad f\left(\left(\begin{array}{r} -2 \\ 2 \end{array}\right)\right) = \left(\begin{array}{r} -2 \\ 2 \\ 0 \\ -2 \end{array}\right)$$

we can still linearly extend f since

$$\begin{aligned}
\phi\left(\left(\begin{array}{r} -2 \\ 2 \end{array}\right)\right) &= \phi\left(-2\left(\begin{array}{c} 1 \\ 0 \end{array}\right) + 2\left(\begin{array}{c} 0 \\ 1 \end{array}\right)\right) \\
&= -2\phi\left(\left(\begin{array}{c} 1 \\ 0 \end{array}\right)\right) + 2\phi\left(\left(\begin{array}{c} 0 \\ 1 \end{array}\right)\right) \\
&= -2\left(\begin{array}{c} 2 \\ 0 \\ 1 \\ 2 \end{array}\right) + 2\left(\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array}\right) = \left(\begin{array}{r} -2 \\ 2 \\ 0 \\ -2 \end{array}\right) \\
&= f\left(\left(\begin{array}{r} -2 \\ 2 \end{array}\right)\right)
\end{aligned}$$

However, if the following was given:

$$f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 2 \end{pmatrix} \qquad f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad f\left(\begin{pmatrix} -2 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

then

$$\begin{aligned}
\phi\left(\begin{pmatrix} -2 \\ 2 \end{pmatrix}\right) &= \phi\left(-2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\
&= -2\phi\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + 2\phi\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\
&= -2\begin{pmatrix} 2 \\ 0 \\ 1 \\ 2 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 0 \\ -2 \end{pmatrix} \\
&\neq f\left(\begin{pmatrix} -2 \\ 2 \end{pmatrix}\right)
\end{aligned}$$

and the function cannot be linearly extended.

3.5.2 $\mathcal{M}_{3\times 2} \to \mathbf{P}_6$ example

Let: $f : \mathcal{M}_{3\times 2} \to \mathbf{P}_6$

$$\begin{aligned}
f\left(\begin{pmatrix} 0 & -5 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}\right) &= 2\,x^5 + 2\,x^4 + 18\,x^3 + 14\,x^2 + 4\,x - 2 \\
f\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\right) &= -5\,x^3 - x^2 + 1 \\
f\left(\begin{pmatrix} 0 & 0 \\ 0 & -1 \\ 5 & 0 \end{pmatrix}\right) &= 39\,x^5 + 15\,x^4 - 36\,x^3 + 74\,x^2 + 38\,x + 16 \\
f\left(\begin{pmatrix} 0 & 4 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}\right) &= -2\,x^5 + x^4 - 24\,x^3 + x + 6 \\
f\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}\right) &= 6\,x^5 + 3\,x^4 - 5\,x^3 + 15\,x^2 + 7\,x + 3 \\
f\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}\right) &= 3\,x^5 + 3\,x^4 - 3\,x^3 + 15\,x^2 + 6\,x + 3
\end{aligned}$$

$$\begin{aligned}
\phi\left(\begin{pmatrix} u_0 & u_1 \\ u_2 & u_3 \\ u_4 & u_5 \end{pmatrix}\right) &= -(3\,u_0 + 2\,u_2 + 9\,u_3 - 6\,u_4 - 2\,u_5)x^5 + (u_2 + 3\,u_4 + 2\,u_5)x^4 \\
&\quad +(2\,u_0 - 5\,u_1 - 4\,u_2 + 11\,u_3 - 5\,u_4 - 7\,u_5)x^3 \\
&\quad -(u_1 - 4\,u_2 - u_3 - 15\,u_4 - 9\,u_5)x^2 \\
&\quad -(u_0 - u_2 + 3\,u_3 - 7\,u_4 - 4\,u_5)x + u_1 + 2\,u_2 - u_3 + 3\,u_4 + 3\,u_5
\end{aligned}$$

$$\phi\left(\begin{pmatrix} -5 & -3 \\ 0 & -3 \\ 8 & 0 \end{pmatrix}\right) = 90\,x^5 + 24\,x^4 - 68\,x^3 + 120\,x^2 + 70\,x + 24$$

$$\phi\left(\begin{pmatrix} -1 & -1 \\ 3 & -2 \\ 5 & 2 \end{pmatrix}\right) = 49\,x^5 + 22\,x^4 - 70\,x^3 + 104\,x^2 + 53\,x + 28$$

3.5.3 $\mathrm{P}_3 \to \mathcal{M}_{2\times 2}$ example

Let

$$\begin{aligned}
f\left(x^2 - 1\right) &= \begin{pmatrix} 5 & -5 \\ 0 & 20 \end{pmatrix} \\
f\left(-4\,x^2 - x + 4\right) &= \begin{pmatrix} -19 & 19 \\ 0 & -76 \end{pmatrix} \\
f\left(1\right) &= \begin{pmatrix} 3 & 0 \\ 1 & -4 \end{pmatrix}
\end{aligned}$$

Extending linearly we have

$$\begin{aligned}
\phi\left(u_2 x^2 + u_1 x + u_0\right) &= \begin{pmatrix} 3\,u_0 - u_1 + 8\,u_2 & u_1 - 5\,u_2 \\ u_0 + u_2 & -4\,u_0 - 4\,u_1 + 16\,u_2 \end{pmatrix} \\
\phi\left(x^2 - 6\right) &= \begin{pmatrix} -10 & -5 \\ -5 & 40 \end{pmatrix} \\
\phi\left(14\,x^2 + 3\,x - 13\right) &= \begin{pmatrix} 70 & -67 \\ 1 & 264 \end{pmatrix}
\end{aligned}$$

3.6 Rank nullity

Theorem 47. *Under a homomorphism, the image of any subspace of the domain is a subspace of the co-domain.*

Proof. Let $\mathbf{S}$ be a subspace of $\phi : \mathbf{V} \to \mathbf{W}$. Then $\phi(S)$ is non-empty because of the zero vector $\vec{0}_V$. If $\phi(\vec{u})$ and $\phi(\vec{v})$ are in the image of $\phi(S)$ then $\phi(a\vec{u} + b\vec{v})$ is also in the image image of ϕ since $a\vec{u} + b\vec{b} \in S$ by the closure of $\mathbf{S}$ $\square$

Definition 48. *The* range space *of a homomorphism $\phi : \mathbf{V} \to \mathbf{W}$ is defined as*

$$\mathcal{R}(\phi) = \{\phi(\vec{v}) \mid \vec{v} \in \mathbf{V}\}$$

The dimension of $\mathcal{R}(\phi)$ is called rank *of ϕ.*

Example: for the linear extension ϕ in Section 3.5.3.

$$\begin{aligned} \mathcal{R}(\phi) &= \left\langle \begin{pmatrix} 5 & -5 \\ 0 & 20 \end{pmatrix} \begin{pmatrix} -19 & 19 \\ 0 & -76 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 1 & -4 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} 5 & -5 \\ 0 & 20 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 1 & -4 \end{pmatrix} \right\rangle \end{aligned}$$

its rank is two.

Examples:

1. derivative transformation $\{1, x, x^2\}$ from $\mathbf{P}_3$ to $\mathbf{P}_3$ has rank two as the image of the derivative is the set of all linear polynomials.

2. derivative transformation $\langle\{\sin x, \cos x\}\rangle$ has rank two as the image of the transformation is all of the span of the two functions.

3. derivative $\{x, \sin x, \cos x\}$ image is $\{1, \sin x, \cos x\}$ the rank is three as the image of the transformation is all of the span of $\{1, \sin x, \cos x\}$

4. Example in Section 3.2.5 has rank three.

Theorem 48. *For any homomorphism the inverse image of a subspace of the co-domain is a subspace of the domain.*

Proof. Let $\phi : \mathbf{V} \to \mathbf{W}$ be a homomorphism and let $\mathbf{T}$ be a subspace of the range space of ϕ. Consider the inverse image of $\mathbf{T}$. It is nonempty because it contains $\vec{0}_V$, since $\phi(\vec{0}_V) = \vec{0}_W$ and $\vec{0}_W$ is an element of $\mathbf{T}$ as $\mathbf{T}$ is a subspace. To finish we show that $\phi^{-1}(\mathbf{T})$ is closed under linear combinations. Let $\vec{v}_1$ and $\vec{v}_2$ be two of its elements, so that $\phi(\vec{v}_1)$ and $\phi(\vec{v}_2)$ are elements of $\mathbf{T}$. Then $c_1\vec{v}_1 + c_2\vec{v}_2$ is an element of the inverse image $\phi^{-1}(\mathbf{T})$ because $\phi(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1\phi(\vec{v}_1) + c_2\phi(\vec{v}_2)$ is a member of $\mathbf{T}$. $\square$

Example: for the linear extension ϕ in Section 3.5.3.

$$\begin{aligned}
\left\langle x^2-1, -4\,x^2-x+4\right\rangle &= \phi^{-1}\left(\left\langle\left(\begin{array}{cc} 5 & -5 \\ 0 & 20 \end{array}\right)\right\rangle\right) \\
\langle 1\rangle &= \phi^{-1}\left(\left\langle\left(\begin{array}{cc} 3 & 0 \\ 1 & -4 \end{array}\right)\right\rangle\right) \\
\mathbf{P}_3 &= \phi^{-1}\left(\left\langle\left(\begin{array}{cc} 5 & -5 \\ 0 & 20 \end{array}\right)\left(\begin{array}{cc} 3 & 0 \\ 1 & -4 \end{array}\right)\right\rangle\right)
\end{aligned}$$

Definition 49 (kernel, null space)**.** *The* null space *or* kernel *of* $\phi : \mathbf{V} \to \mathbf{W}$ *is*

$$\ker(\phi) = \{\vec{v} \in \mathbf{V} \mid \phi(\vec{v}) = \vec{0}_W\}.$$

The dimension of $\ker(\phi)$ *is called* nullity *of* ϕ*.*

The kernel may be denoted as $\phi^{-1}\left(\vec{0}\right)$ or $\mathcal{N}(\phi)$.

Example for the linear extension ϕ in Section 3.5.3.

$$\begin{aligned}
\ker(\phi) &= \phi^{-1}\left(\left\langle\left(\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right)\right\rangle\right) \\
&= \left\langle -x^2-5\,x+1\right\rangle
\end{aligned}$$

it has nullity one.

Example the null space of the derivative operator is the set of all constants and it has nullity one.

Theorem 49. *Let* $\phi : \mathbf{V} \to \mathbf{W}$ *be a linear map. Then* $\dim \mathbf{V}$ *equals the sum of the nullity of* ϕ *plus the rank of* ϕ

$$\dim \mathbf{V} = \dim \mathcal{R}(\phi) + \dim \mathcal{N}(\phi)$$

Proof. Let $\phi : \mathbf{V} \to \mathbf{W}$ be linear and let $B_N = \{\vec{\beta}_1, \ldots, \vec{\beta}_k\}$ be a basis for the null space. Expand that to a basis $B_V = \{\vec{\beta}_1, \ldots, \vec{\beta}_k, \vec{\beta}_{k+1}, \ldots, \vec{\beta}_n\}$ for the entire domain. We shall show that $B_R = \{\phi(\vec{\beta}_{k+1}), \ldots, \phi(\vec{\beta}_n)\}$ is a basis for the range space. Then counting the size of the bases gives the result.

To see that B_R is linearly independent, consider

$$\vec{0}_W = c_{k+1}\phi(\vec{\beta}_{k+1}) + \cdots + c_n\phi(\vec{\beta}_n) = \phi(c_{k+1}\vec{\beta}_{k+1} + \cdots + c_n\vec{\beta}_n)$$

and so $c_{k+1}\vec{\beta}_{k+1} + \cdots + c_n\vec{\beta}_n \in \mathcal{N}\phi$. As B_N is a basis for the null space there are scalars $c_1, \ldots, c_k$ satisfying this relationship.

$$c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k = c_{k+1}\vec{\beta}_{k+1} + \cdots + c_n\vec{\beta}_n$$

equivalently

$$c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k - c_{k+1}\vec{\beta}_{k+1} - \cdots - c_n\vec{\beta}_n = \vec{0}_V$$

But this is an equation among members of B_V, which is a basis for V, so each c_i equals 0. Therefore B_R is linearly independent.

To show that B_R spans the range space consider a member of the range space $\phi(\vec{v})$. Express $\vec{v}$ as a linear combination $\vec{v} = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n$ of members of B_V. This gives

$$\begin{aligned}
\phi(\vec{v}) &= \phi(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n + c_{k+1}(\vec{\beta}_{k+1}) + \cdots + c_n(\vec{\beta}_n)) \\
&= c_1\phi(\vec{\beta}_1) + \cdots + c_k\phi(\vec{\beta}_k) + c_{k+1}\phi(\vec{\beta}_{k+1}) + \cdots + c_n\phi(\vec{\beta}_n) \\
&= c_1\vec{0}_W + \cdots + c_k\vec{0}_W + c_{k+1}\phi(\vec{\beta}_{k+1}) + \cdots + c_n\phi(\vec{\beta}_n) \\
&= c_{k+1}\phi(\vec{\beta}_{k+1}) + \cdots + c_n\phi(\vec{\beta}_n)
\end{aligned}$$

and since $\vec{\beta}_1, \ldots, \vec{\beta}_k$ are in the null space, we have that $h(\vec{v}) = \vec{0}_W + \cdots + \vec{0}_W + c_{k+1}h(\vec{\beta}_{k+1}) + \cdots + c_nh(\vec{\beta}_n)$. Thus, $h(\vec{v})$ is a linear combination of members of B_R, and so B_R spans the range space. □

3.7 Matrix representation of linear maps

Question: How to describe compactly linear transformations?

http://dev.w3.org/csswg/css3-3d-transforms/

> matrix3d(<number>,<number>,<number>,<number>,<number>,<number>,
> <number>,<number>,<number>,<number>,<number>,<number>,<number>,
> <number>,<number>,<number>)
>
> specifies a 3D transformation as a 4x4 homogeneous matrix of 16 values in column-major order.

Definition 50. *Suppose that* $\mathbf{U}$ *and* $\mathbf{W}$ *are vector spaces of dimensions* n *and* m *with bases* $B = \left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right)$ *and* $E = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m)$*, and that* $\phi : \mathbf{U} \to \mathbf{W}$ *is a linear map. If*

$$\mathcal{R}_E(\phi(\vec{b}_1)) = \begin{pmatrix} h_{1,1} \\ h_{2,1} \\ \vdots \\ h_{m,1} \end{pmatrix}_E \quad \ldots \quad \mathcal{R}_E(\phi(\vec{b}_n)) = \begin{pmatrix} h_{1,n} \\ h_{2,n} \\ \vdots \\ h_{m,n} \end{pmatrix}_E$$

then

$$\begin{aligned}
\mathcal{R}_{B\to E}(\phi) &= \begin{pmatrix} | & | & & | \\ \mathcal{R}_E\left(\phi(\vec{b}_1)\right) & \mathcal{R}_E\left(\phi(\vec{b}_2)\right) & \dots & \mathcal{R}_E\left(\phi(\vec{b}_n)\right) \\ | & | & & | \end{pmatrix} \\
&= \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,n} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \cdots & h_{m,n} \end{pmatrix}_{B\to E}
\end{aligned}$$

is the matrix representation of ϕ with respect to $B \to E$.

3.7.1 $\mathrm{d} : \mathbf{P}_3 \to \mathbf{P}_2$

Let $\mathrm{d} : \mathbf{P}_3 \to \mathbf{P}_2$ denote the derivative function, where $\mathbf{P}_3$ the vector space of polynomials of degree at most three and $\mathbf{P}_2$ is the vector space of polynomials of degree at most two. For basis given below find the matrix representation of d from basis B_{P_3} to basis B_{P_2} i.e.,

$$\mathcal{R}_{B_{P_3}\to B_{P_2}}(\mathrm{d})$$

Bases $B \to E$

Given $\mathbf{P}_3 = \langle B \rangle = \left\langle \vec{b}_0, \vec{b}_1, \vec{b}_2, \vec{b}_3 \right\rangle$ where

$$\begin{aligned}
\vec{b}_0 &= 1 \\
\vec{b}_1 &= x \\
\vec{b}_2 &= x^2 \\
\vec{b}_3 &= x^3
\end{aligned}$$

and $\mathbf{P}_2 = \langle E \rangle = \langle \vec{e}_0, \vec{e}_1, \vec{e}_2 \rangle$ where

$$\begin{aligned}
\vec{e}_0 &= 1 \\
\vec{e}_1 &= x \\
\vec{e}_2 &= x^2
\end{aligned}$$

For basis vector $\vec{b}_0$, we compute

$$\begin{aligned}
\mathrm{d}\,(1) &= 0 \\
&= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\
&= 0\vec{e}_0 + 0\vec{e}_1 + 0\vec{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}_E
\end{aligned}$$

where the coordinates are obtained by solving

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{c} x_0 \\ x_1 \\ x_2 \end{array}\right)=\left(\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right)$$

For basis vector $\vec{b}_1$, we compute

$$\begin{aligned} \mathrm{d}(x) &= 1 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 1\vec{e}_0 + 0\vec{e}_1 + 0\vec{e}_2 = \left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right)_E \end{aligned}$$

where the coordinates are obtained by solving

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{c} x_0 \\ x_1 \\ x_2 \end{array}\right)=\left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right)$$

For basis vector $\vec{b}_2$, we compute

$$\begin{aligned} \mathrm{d}\left(x^2\right) &= 2\,x \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 0\vec{e}_0 + 2\vec{e}_1 + 0\vec{e}_2 = \left(\begin{array}{c} 0 \\ 2 \\ 0 \end{array}\right)_E \end{aligned}$$

where the coordinates are obtained by solving

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{c} x_0 \\ x_1 \\ x_2 \end{array}\right)=\left(\begin{array}{c} 0 \\ 2 \\ 0 \end{array}\right)$$

For basis vector $\vec{b}_3$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3\right) &= 3\,x^2 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 0\vec{e}_0 + 0\vec{e}_1 + 3\vec{e}_2 = \left(\begin{array}{c} 0 \\ 0 \\ 3 \end{array}\right)_E \end{aligned}$$

where the coordinates are obtained by solving

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{c} x_0 \\ x_1 \\ x_2 \end{array}\right)=\left(\begin{array}{c} 0 \\ 0 \\ 3 \end{array}\right)$$

Solution is

$$\mathcal{R}_{B\to E}(\mathrm{d}) = \left(\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{array}\right)$$

Bases $B \to G$

Given $\mathbf{P}_3 = \langle B \rangle = \left\langle \vec{b}_0, \vec{b}_1, \vec{b}_2, \vec{b}_3 \right\rangle$ where

$$\begin{array}{rcl} \vec{b}_0 & = & 1 \\ \vec{b}_1 & = & x \\ \vec{b}_2 & = & x^2 \\ \vec{b}_3 & = & x^3 \end{array}$$

and $\mathbf{P}_2 = \langle G \rangle = \langle \vec{g}_1, \vec{g}_2, \vec{g}_3 \rangle$ where

$$\begin{array}{rcl} \vec{g}_0 & = & x^2 + x + 1 \\ \vec{g}_1 & = & x^2 + x \\ \vec{g}_2 & = & x^2 \end{array}$$

For basis vector $\vec{b}_0$, we compute

$$\begin{array}{rcl} \mathrm{d}\,(1) & = & 0 \\ & = & x_0\vec{g}_0 + x_1\vec{g}_1 + x_2\vec{g}_2 \\ & = & 0\vec{g}_0 + 0\vec{g}_1 + 0\vec{g}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}_G \end{array}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_1$, we compute

$$\begin{array}{rcl} \mathrm{d}\,(x) & = & 1 \\ & = & x_0\vec{g}_0 + x_1\vec{g}_1 + x_2\vec{g}_2 \\ & = & 1\vec{g}_0 - 1\vec{g}_1 + 0\vec{g}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}_G \end{array}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_2$, we compute

$$\begin{aligned} \mathrm{d}\left(x^2\right) &= 2\,x \\ &= x_0\vec{g}_0 + x_1\vec{g}_1 + x_2\vec{g}_2 \\ &= 0\vec{g}_0 + 2\vec{g}_1 - 2\vec{g}_2 = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}_G \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_3$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3\right) &= 3\,x^2 \\ &= x_0\vec{g}_0 + x_1\vec{g}_1 + x_2\vec{g}_2 \\ &= 0\vec{g}_0 + 0\vec{g}_1 + 3\vec{g}_2 = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}_G \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{B\to G}\left(\mathrm{d}\right) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -2 & 3 \end{pmatrix}$$

Bases $A \to E$

Given $\mathbf{P}_3 = \langle A \rangle = \langle \vec{a}_0, \vec{a}_1, \vec{a}_2, \vec{a}_3 \rangle$ where

$$\begin{aligned} \vec{a}_0 &= -x^3 + 2\,x^2 + 2\,x + 1 \\ \vec{a}_1 &= x^3 + x^2 - 2 \\ \vec{a}_2 &= x^3 + 3\,x^2 + 3\,x + 2 \\ \vec{a}_3 &= x^3 + x^2 + x + 1 \end{aligned}$$

and $\mathbf{P}_2 = \langle E \rangle = \langle \vec{e}_0, \vec{e}_1, \vec{e}_2 \rangle$ where

$$\begin{aligned} \vec{e}_0 &= 1 \\ \vec{e}_1 &= x \\ \vec{e}_2 &= x^2 \end{aligned}$$

For basis vector $\vec{a}_0$, we compute

$$\begin{aligned} \mathrm{d}\left(-x^3 + 2\,x^2 + 2\,x + 1\right) &= -3\,x^2 + 4\,x + 2 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 2\vec{e}_0 + 4\vec{e}_1 - 3\vec{e}_2 = \begin{pmatrix} 2 \\ 4 \\ -3 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -3 \end{pmatrix}$$

For basis vector $\vec{a}_1$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3 + x^2 - 2\right) &= 3\,x^2 + 2\,x \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 0\vec{e}_0 + 2\vec{e}_1 + 3\vec{e}_2 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}$$

For basis vector $\vec{a}_2$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3 + 3\,x^2 + 3\,x + 2\right) &= 3\,x^2 + 6\,x + 3 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 3\vec{e}_0 + 6\vec{e}_1 + 3\vec{e}_2 = \begin{pmatrix} 3 \\ 6 \\ 3 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 3 \end{pmatrix}$$

For basis vector $\vec{a}_3$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3 + x^2 + x + 1\right) &= 3\,x^2 + 2\,x + 1 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 1\vec{e}_0 + 2\vec{e}_1 + 3\vec{e}_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{A\to E}\left(\mathrm{d}\right) = \begin{pmatrix} 2 & 0 & 3 & 1 \\ 4 & 2 & 6 & 2 \\ -3 & 3 & 3 & 3 \end{pmatrix}$$

Bases $A \to H$

Given $\mathbf{P}_3 = \langle A \rangle = \langle \vec{a}_0, \vec{a}_1, \vec{a}_2, \vec{a}_3 \rangle$ where

$$\begin{aligned} \vec{a}_0 &= -x^3 + 2\,x^2 + 2\,x + 1 \\ \vec{a}_1 &= x^3 + x^2 - 2 \\ \vec{a}_2 &= x^3 + 3\,x^2 + 3\,x + 2 \\ \vec{a}_3 &= x^3 + x^2 + x + 1 \end{aligned}$$

and $\mathbf{P}_2 = \langle H \rangle = \left\langle \vec{h}_0, \vec{h}_1, \vec{h}_2 \right\rangle$ where

$$\begin{aligned} \vec{h}_0 &= 3\,x^2 + 2\,x + 4 \\ \vec{h}_1 &= x^2 + 2\,x + 3 \\ \vec{h}_2 &= x + 1 \end{aligned}$$

For basis vector $\vec{a}_0$, we compute

$$\begin{aligned} \mathrm{d}\left(-x^3 + 2\,x^2 + 2\,x + 1\right) &= -3\,x^2 + 4\,x + 2 \\ &= x_0\vec{h}_0 + x_1\vec{h}_1 + x_2\vec{h}_2 \\ &= -1\vec{h}_0 + 0\vec{h}_1 + 6\vec{h}_2 = \begin{pmatrix} -1 \\ 0 \\ 6 \end{pmatrix}_H \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -3 \end{pmatrix}$$

For basis vector $\vec{a}_1$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3 + x^2 - 2\right) &= 3\,x^2 + 2\,x \\ &= x_0\vec{h}_0 + x_1\vec{h}_1 + x_2\vec{h}_2 \\ &= 5\vec{h}_0 - 12\vec{h}_1 + 16\vec{h}_2 = \begin{pmatrix} 5 \\ -12 \\ 16 \end{pmatrix}_H \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}$$

For basis vector $\vec{a}_2$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3 + 3\,x^2 + 3\,x + 2\right) &= 3\,x^2 + 6\,x + 3 \\ &= x_0\vec{h}_0 + x_1\vec{h}_1 + x_2\vec{h}_2 \\ &= 6\vec{h}_0 - 15\vec{h}_1 + 24\vec{h}_2 = \begin{pmatrix} 6 \\ -15 \\ 24 \end{pmatrix}_H \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 3 \end{pmatrix}$$

For basis vector $\vec{a}_3$, we compute

$$\begin{aligned} \mathrm{d}\left(x^3 + x^2 + x + 1\right) &= 3\,x^2 + 2\,x + 1 \\ &= x_0\vec{h}_0 + x_1\vec{h}_1 + x_2\vec{h}_2 \\ &= 4\vec{h}_0 - 9\vec{h}_1 + 12\vec{h}_2 = \begin{pmatrix} 4 \\ -9 \\ 12 \end{pmatrix}_H \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{A\to H}\left(\mathrm{d}\right) = \begin{pmatrix} -1 & 5 & 6 & 4 \\ 0 & -12 & -15 & -9 \\ 6 & 16 & 24 & 12 \end{pmatrix}$$

Combining it together

Consider now the polynomial and its representation is various basis

$$
\begin{aligned}
\vec{p} &= p(x) = -x^3 + 3x^2 + 1 \\
\mathcal{R}_B(\vec{p}) &= \begin{pmatrix} 1 \\ 0 \\ 3 \\ -1 \end{pmatrix}_B \\
\mathcal{R}_A(\vec{p}) &= \begin{pmatrix} 18 \\ 3 \\ -25 \\ 39 \end{pmatrix}_A
\end{aligned}
$$

For its derivative

$$
\begin{aligned}
\vec{q} &= d(p(x)) = -3x^2 + 6x \\
\mathcal{R}_E(d(\vec{p})) &= \begin{pmatrix} 0 \\ 6 \\ -3 \end{pmatrix}_E \\
\mathcal{R}_G(d(\vec{p})) &= \begin{pmatrix} 0 \\ 6 \\ -9 \end{pmatrix}_G \\
\mathcal{R}_H(d(\vec{p})) &= \begin{pmatrix} 3 \\ -12 \\ 24 \end{pmatrix}_H
\end{aligned}
$$

Then

$$\mathcal{R}_E\left(d(\vec{p})\right) = \mathcal{R}_{B\to E}(d)\mathcal{R}_B\left(\vec{p}\right) \quad\to\quad \begin{pmatrix} 0 \\ 6 \\ -3 \end{pmatrix}_E = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}_{B\to E} \begin{pmatrix} 1 \\ 0 \\ 3 \\ -1 \end{pmatrix}_B$$

$$\mathcal{R}_G\left(d(\vec{p})\right) = \mathcal{R}_{B\to G}(d)\mathcal{R}_B\left(\vec{p}\right) \quad\to\quad \begin{pmatrix} 0 \\ 6 \\ -9 \end{pmatrix}_G = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -2 & 3 \end{pmatrix}_{B\to G} \begin{pmatrix} 1 \\ 0 \\ 3 \\ -1 \end{pmatrix}_B$$

$$\mathcal{R}_E\left(d(\vec{p})\right) = \mathcal{R}_{A\to E}(d)\mathcal{R}_A\left(\vec{p}\right) \quad\to\quad \begin{pmatrix} 0 \\ 6 \\ -3 \end{pmatrix}_E = \begin{pmatrix} 2 & 0 & 3 & 1 \\ 4 & 2 & 6 & 2 \\ -3 & 3 & 3 & 3 \end{pmatrix}_{A\to E} \begin{pmatrix} 18 \\ 3 \\ -25 \\ 39 \end{pmatrix}_A$$

$$\mathcal{R}_H\left(d(\vec{p})\right) = \mathcal{R}_{A\to H}(d)\mathcal{R}_A\left(\vec{p}\right) \quad\to\quad \begin{pmatrix} 3 \\ -12 \\ 24 \end{pmatrix}_H = \begin{pmatrix} -1 & 5 & 6 & 4 \\ 0 & -12 & -15 & -9 \\ 6 & 16 & 24 & 12 \end{pmatrix}_{A\to H} \begin{pmatrix} 18 \\ 3 \\ -25 \\ 39 \end{pmatrix}_A$$

3.7.2 $M_{2\times 2} \to \mathbf{P}_2$

Problem: Let $\phi : M_{2\times 2} \to \mathbf{P}_2$ be defined as

$$\phi\left(\begin{pmatrix} u_0 & u_1 \\ u_2 & u_3 \end{pmatrix}\right) = -(3\,u_0 - 6\,u_1 - u_2 + 3\,u_3)x^2 - (u_0 - 2\,u_1 - u_2 + u_3)x - u_2$$

Given $M_{2\times 2} = \langle B \rangle$

$$\begin{aligned} \vec{b}_0 &= \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \\ \vec{b}_1 &= \begin{pmatrix} -1 & 1 \\ 2 & 0 \end{pmatrix} \\ \vec{b}_2 &= \begin{pmatrix} 0 & -3 \\ -6 & -1 \end{pmatrix} \\ \vec{b}_3 &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Given $\mathbf{P}_2 = \langle D \rangle$

$$\begin{aligned} \vec{d}_0 &= -2\,x^2 + 1 \\ \vec{d}_1 &= 2\,x^2 - x - 1 \\ \vec{d}_2 &= -x^2 + 1 \end{aligned}$$

Find the matrix representation of ϕ from basis B to basis D i.e.,

$$\mathcal{R}_{B\to D}(\phi)$$

Solution: For basis vector $\vec{b}_0$, we compute

$$\begin{aligned}
\phi\left(\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}\right) &= 8\,x^2 + 4\,x - 2 \\
&= x_0\vec{d}_0 + x_1\vec{d}_1 + x_2\vec{d}_2 \\
&= -10\vec{d}_0 - 4\vec{d}_1 + 4\vec{d}_2 = \begin{pmatrix} -10 \\ -4 \\ 4 \end{pmatrix}_D
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ -2 & 2 & -1 \end{pmatrix}\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \\ 8 \end{pmatrix}$$

For basis vector $\vec{b}_1$, we compute

$$\begin{aligned}
\phi\left(\begin{pmatrix} -1 & 1 \\ 2 & 0 \end{pmatrix}\right) &= 11\,x^2 + 5\,x - 2 \\
&= x_0\vec{d}_0 + x_1\vec{d}_1 + x_2\vec{d}_2 \\
&= -14\vec{d}_0 - 5\vec{d}_1 + 7\vec{d}_2 = \begin{pmatrix} -14 \\ -5 \\ 7 \end{pmatrix}_D
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ -2 & 2 & -1 \end{pmatrix}\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 5 \\ 11 \end{pmatrix}$$

For basis vector $\vec{b}_2$, we compute

$$\begin{aligned}
\phi\left(\begin{pmatrix} 0 & -3 \\ -6 & -1 \end{pmatrix}\right) &= -21\,x^2 - 11\,x + 6 \\
&= x_0\vec{d}_0 + x_1\vec{d}_1 + x_2\vec{d}_2 \\
&= 26\vec{d}_0 + 11\vec{d}_1 - 9\vec{d}_2 = \begin{pmatrix} 26 \\ 11 \\ -9 \end{pmatrix}_D
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ -2 & 2 & -1 \end{pmatrix}\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ -11 \\ -21 \end{pmatrix}$$

For basis vector $\vec{b}_3$, we compute

$$\begin{aligned}
\phi\left(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right) &= x^2 + x - 1 \\
&= x_0\vec{d}_0 + x_1\vec{d}_1 + x_2\vec{d}_2 \\
&= -1\vec{d}_0 - 1\vec{d}_1 - 1\vec{d}_2 = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}_D
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ -2 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Combining together the representation is

$$\mathcal{R}_{B\to D}(\phi) = \begin{pmatrix} -10 & -14 & 26 & -1 \\ -4 & -5 & 11 & -1 \\ 4 & 7 & -9 & -1 \end{pmatrix}$$

3.7.3 $\mathbf{P}_2 \to \mathcal{D}_2$

Problem: Consider $\phi(\cdot) : \mathbf{P}_2 \to \mathcal{D}_2$ with domain polynomials of degree at most two $\mathbf{P}_2$ and range 2×2 diagonal matrices $\mathcal{D}_2$ defined as:

$$\phi\left(u_2x^2 + u_1x + u_0\right) = \begin{pmatrix} u_0 - 3\,u_2 & 0 \\ 0 & -u_1 \end{pmatrix}$$

For basis B

$$\begin{aligned}
\vec{b}_0 &= 3\,x^2 + x \\
\vec{b}_1 &= -8\,x^2 - 3\,x \\
\vec{b}_2 &= 1
\end{aligned}$$

and basis D

$$\vec{d}_0 = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \vec{d}_1 = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$$

Find the matrix representation of ϕ from basis B to basis D i.e.,

$$\mathcal{R}_{B\to D}(\phi)$$

Solution: For basis vector $\vec{a}_0$, we compute

$$\begin{aligned}
\phi\left(3\,x^2 + x\right) &= \begin{pmatrix} -9 & 0 \\ 0 & -1 \end{pmatrix} \\
&= x_0\vec{d}_0 + x_1\vec{d}_1 \\
&= 11\vec{d}_0 + 1\vec{d}_1 = \begin{pmatrix} 11 \\ 1 \end{pmatrix}_D
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} -9 \\ -1 \end{pmatrix}$$

For basis vector $\vec{a}_1$, we compute

$$\begin{aligned}
\phi\left(-8\,x^2 - 3\,x\right) &= \begin{pmatrix} 24 & 0 \\ 0 & 3 \end{pmatrix} \\
&= x_0\vec{d}_0 + x_1\vec{d}_1 \\
&= -30\vec{d}_0 - 3\vec{d}_1 = \begin{pmatrix} -30 \\ -3 \end{pmatrix}_D
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 24 \\ 3 \end{pmatrix}$$

For basis vector $\vec{a}_2$, we compute

$$\begin{aligned}
\phi\left(1\right) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\
&= x_0\vec{d}_0 + x_1\vec{d}_1 \\
&= -1\vec{d}_0 + 0\vec{d}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}_D
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{B\to D}(\phi) = \begin{pmatrix} 11 & -30 & -1 \\ 1 & -3 & 0 \end{pmatrix}$$

3.7.4 $\mathcal{M}_{2\times 3} \to \mathcal{M}_{2\times 3}$

The following function is a variant of the linear map from §3.2.5

$$\phi : \mathcal{M}_{2\times 3} \to \mathcal{M}_{2\times 3}$$

defined as

$$\phi\left(\begin{pmatrix} u_0 & u_1 \\ u_2 & u_3 \\ u_4 & u_5 \end{pmatrix}\right) = \begin{pmatrix} u_0 + u_2 + u_5 & u_1 - 2\,u_3 \\ u_0 - 3\,u_1 + 2\,u_3 & 2\,u_1 + u_5 \\ u_1 + u_2 - 2\,u_3 & 2\,u_1 - 4\,u_3 \end{pmatrix}$$

Problem: for basis B

$$\begin{aligned}
\vec{b}_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\vec{b}_1 &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\vec{b}_2 &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \\
\vec{b}_3 &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\
\vec{b}_4 &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \\
\vec{b}_5 &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

find the matrix representation of ϕ from basis B to the same basis B.

Solution: For basis vector $\vec{b}_0$, we compute

$$\begin{aligned}
\phi\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \\
&= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 + x_4\vec{b}_4 + x_5\vec{b}_5 \\
&= 1\vec{b}_0 + 0\vec{b}_1 + 1\vec{b}_2 + 0\vec{b}_3 + 0\vec{b}_4 + 0\vec{b}_5 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}_B
\end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_1$, we compute

$$\begin{aligned} \phi\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 0 & 1 \\ -3 & 2 \\ 1 & 2 \end{pmatrix} \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 + x_4\vec{b}_4 + x_5\vec{b}_5 \\ &= 0\vec{b}_0 + 1\vec{b}_1 - 3\vec{b}_2 + 2\vec{b}_3 + 1\vec{b}_4 + 2\vec{b}_5 = \begin{pmatrix} 0 \\ 1 \\ -3 \\ 2 \\ 1 \\ 2 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -3 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

For basis vector $\vec{b}_2$, we compute

$$\begin{aligned} \phi\left(\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 + x_4\vec{b}_4 + x_5\vec{b}_5 \\ &= 1\vec{b}_0 + 0\vec{b}_1 + 0\vec{b}_2 + 0\vec{b}_3 + 1\vec{b}_4 + 0\vec{b}_5 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_3$, we compute

$$\begin{aligned} \phi\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 0 & -2 \\ 2 & 0 \\ -2 & -4 \end{pmatrix} \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 + x_4\vec{b}_4 + x_5\vec{b}_5 \\ &= 0\vec{b}_0 - 2\vec{b}_1 + 2\vec{b}_2 + 0\vec{b}_3 - 2\vec{b}_4 - 4\vec{b}_5 = \begin{pmatrix} 0 \\ -2 \\ 2 \\ 0 \\ -2 \\ -4 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 2 \\ 0 \\ -2 \\ -4 \end{pmatrix}$$

For basis vector $\vec{b}_4$, we compute

$$\begin{aligned} \phi\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 + x_4\vec{b}_4 + x_5\vec{b}_5 \\ &= 0\vec{b}_0 + 0\vec{b}_1 + 0\vec{b}_2 + 0\vec{b}_3 + 0\vec{b}_4 + 0\vec{b}_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_5$, we compute

$$\begin{aligned} \phi\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}\right) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\ &= x_0\vec{d}_0 + x_1\vec{d}_1 + x_2\vec{d}_2 + x_3\vec{d}_3 + x_4\vec{d}_4 + x_5\vec{d}_5 \\ &= 1\vec{d}_0 + 0\vec{d}_1 + 0\vec{d}_2 + 1\vec{d}_3 + 0\vec{d}_4 + 0\vec{d}_5 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{B\to B}(\phi) = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 & 0 & 0 \\ 1 & -3 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -2 & 0 & 0 \\ 0 & 2 & 0 & -4 & 0 & 0 \end{pmatrix}$$

3.8 Change of basis

Definition 51. *The* change of basis matrix *for bases $B, A \subset \mathbf{V}$ is the representation of the identity map $id : \mathbf{V} \to \mathbf{V}$ with respect to those bases.*

$$\mathcal{R}_{B\to A}(\phi) = \begin{pmatrix} | & | & & | \\ \mathcal{R}_A(\vec{b}_1) & \mathcal{R}_A(\vec{b}_2) & \dots & \mathcal{R}_A(\vec{b}_n) \\ | & | & & | \end{pmatrix} = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ & \vdots & & \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{pmatrix}$$

is the matrix representation of ϕ with respect to $B \to A$.

3.8.1 $\mathbf{P}_3 \to \mathbf{P}_3$

Using $\mathbf{P}_3$ basis from Section 3.7.1, given $\mathbf{P}_3 = \langle B \rangle$

$$\begin{aligned} \vec{b}_0 &= 1 \\ \vec{b}_1 &= x \\ \vec{b}_2 &= x^2 \\ \vec{b}_3 &= x^3 \end{aligned}$$

and $\mathbf{P}_3 = \langle A \rangle$

$$\begin{aligned} \vec{a}_0 &= -x^3 + 2\,x^2 + 2\,x + 1 \\ \vec{a}_1 &= x^3 + x^2 - 2 \\ \vec{a}_2 &= x^3 + 3\,x^2 + 3\,x + 2 \\ \vec{a}_3 &= x^3 + x^2 + x + 1 \end{aligned}$$

Bases $A \to B$

For basis vector $\vec{a}_0$, we compute

$$\begin{aligned} id\left(-x^3 + 2\,x^2 + 2\,x + 1\right) &= -x^3 + 2\,x^2 + 2\,x + 1 \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 \\ &= 1\vec{b}_0 + 2\vec{b}_1 + 2\vec{b}_2 - 1\vec{b}_3 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ -1 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \\ -1 \end{pmatrix}$$

For basis vector $\vec{a}_1$, we compute

$$\begin{aligned} id\left(x^3 + x^2 - 2\right) &= x^3 + x^2 - 2 \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 \\ &= -2\vec{b}_0 + 0\vec{b}_1 + 1\vec{b}_2 + 1\vec{b}_3 = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

For basis vector $\vec{a}_2$, we compute

$$\begin{aligned} id\left(x^3 + 3\,x^2 + 3\,x + 2\right) &= x^3 + 3\,x^2 + 3\,x + 2 \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 \\ &= 2\vec{b}_0 + 3\vec{b}_1 + 3\vec{b}_2 + 1\vec{b}_3 = \begin{pmatrix} 2 \\ 3 \\ 3 \\ 1 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 3 \\ 1 \end{pmatrix}$$

For basis vector $\vec{a}_3$, we compute

$$\begin{aligned} id\left(x^3 + x^2 + x + 1\right) &= x^3 + x^2 + x + 1 \\ &= x_0\vec{b}_0 + x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 \\ &= 1\vec{b}_0 + 1\vec{b}_1 + 1\vec{b}_2 + 1\vec{b}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}_B \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{A\to B} = \begin{pmatrix} 1 & -2 & 2 & 1 \\ 2 & 0 & 3 & 1 \\ 2 & 1 & 3 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix}$$

Bases $B \to A$

For basis vector $\vec{b}_0$, we compute

$$\begin{aligned} id\,(1) &= 1 \\ &= x_0\vec{a}_0 + x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 \\ &= 2\vec{a}_0 + 0\vec{a}_1 - 3\vec{a}_2 + 5\vec{a}_3 = \begin{pmatrix} 2 \\ 0 \\ -3 \\ 5 \end{pmatrix}_A \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -2 & 2 & 1 \\ 2 & 0 & 3 & 1 \\ 2 & 1 & 3 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_1$, we compute

$$\begin{aligned} id\,(x) &= x \\ &= x_0\vec{a}_0 + x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 \\ &= -6\vec{a}_0 - 1\vec{a}_1 + 9\vec{a}_2 - 14\vec{a}_3 = \begin{pmatrix} -6 \\ -1 \\ 9 \\ -14 \end{pmatrix}_A \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -2 & 2 & 1 \\ 2 & 0 & 3 & 1 \\ 2 & 1 & 3 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_2$, we compute

$$\begin{aligned} id\left(x^2\right) &= x^2 \\ &= x_0\vec{a}_0 + x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 \\ &= 5\vec{a}_0 + 1\vec{a}_1 - 7\vec{a}_2 + 11\vec{a}_3 = \begin{pmatrix} 5 \\ 1 \\ -7 \\ 11 \end{pmatrix}_A \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -2 & 2 & 1 \\ 2 & 0 & 3 & 1 \\ 2 & 1 & 3 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

For basis vector $\vec{b}_3$, we compute

$$\begin{aligned} id\left(x^3\right) &= x^3 \\ &= x_0\vec{a}_0 + x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 \\ &= -1\vec{a}_0 + 0\vec{a}_1 + 1\vec{a}_2 - 1\vec{a}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}_A \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & -2 & 2 & 1 \\ 2 & 0 & 3 & 1 \\ 2 & 1 & 3 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{B\to A} = \begin{pmatrix} 2 & -6 & 5 & -1 \\ 0 & -1 & 1 & 0 \\ -3 & 9 & -7 & 1 \\ 5 & -14 & 11 & -1 \end{pmatrix}$$

Important observation

$$\mathcal{R}_{B\to A}\mathcal{R}_{A\to B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3.8.2 $\mathbf{P}_2 \to \mathbf{P}_2$

Using $\mathbf{P}_2$ basis from Section 3.7.1, given $\mathbf{P}_2 = \langle E\rangle$

$$\begin{aligned} \vec{e}_0 &= 1 \\ \vec{e}_1 &= x \\ \vec{e}_2 &= x^2 \end{aligned}$$

and $\mathbf{P}_2 = \langle G\rangle$

$$\begin{aligned} \vec{g}_0 &= x^2 + x + 1 \\ \vec{g}_1 &= x^2 + x \\ \vec{g}_2 &= x^2 \end{aligned}$$

and $\mathbf{P}_2 = \langle H \rangle$

$$\begin{array}{rcl} \vec{h}_0 & = & 3\,x^2 + 2\,x + 4 \\ \vec{h}_1 & = & x^2 + 2\,x + 3 \\ \vec{h}_2 & = & x + 1 \end{array}$$

Bases $E \to G$

For basis vector $\vec{e}_0$, we compute

$$\begin{array}{rcl} id\,(1) & = & 1 \\ & = & x_0\vec{g}_0 + x_1\vec{g}_1 + x_2\vec{g}_2 \\ & = & 1\vec{g}_0 - 1\vec{g}_1 + 0\vec{g}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}_G \end{array}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{e}_1$, we compute

$$\begin{array}{rcl} id\,(x) & = & x \\ & = & x_0\vec{g}_0 + x_1\vec{g}_1 + x_2\vec{g}_2 \\ & = & 0\vec{g}_0 + 1\vec{g}_1 - 1\vec{g}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}_G \end{array}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

For basis vector $\vec{e}_2$, we compute

$$\begin{array}{rcl} id\left(x^2\right) & = & x^2 \\ & = & x_0\vec{g}_0 + x_1\vec{g}_1 + x_2\vec{g}_2 \\ & = & 0\vec{g}_0 + 0\vec{g}_1 + 1\vec{g}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_G \end{array}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{E\to G} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

Bases $G \to E$

For basis vector $\vec{g}_0$, we compute

$$\begin{aligned} id\left(x^2+x+1\right) &= x^2+x+1 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 1\vec{e}_0 + 1\vec{e}_1 + 1\vec{e}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

For basis vector $\vec{g}_1$, we compute

$$\begin{aligned} id\left(x^2+x\right) &= x^2+x \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 0\vec{e}_0 + 1\vec{e}_1 + 1\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

For basis vector $\vec{g}_2$, we compute

$$\begin{aligned} id\left(x^2\right) &= x^2 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 0\vec{e}_0 + 0\vec{e}_1 + 1\vec{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{G \to E} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Note that the systems of linear equations are trivial to solve so getting the matrix representation $G \to E$ is trivial. We still have

$$\mathcal{R}_{G \to E} \mathcal{R}_{E \to G} = I$$

Bases $H \to E$

For basis vector $\vec{h}_0$, we compute

$$\begin{aligned} id\left(3\,x^2 + 2\,x + 4\right) &= 3\,x^2 + 2\,x + 4 \\ &= x_0 \vec{e}_0 + x_1 \vec{e}_1 + x_2 \vec{e}_2 \\ &= 4\vec{e}_0 + 2\vec{e}_1 + 3\vec{e}_2 = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}$$

For basis vector $\vec{h}_1$, we compute

$$\begin{aligned} id\left(x^2 + 2\,x + 3\right) &= x^2 + 2\,x + 3 \\ &= x_0 \vec{e}_0 + x_1 \vec{e}_1 + x_2 \vec{e}_2 \\ &= 3\vec{e}_0 + 2\vec{e}_1 + 1\vec{e}_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

For basis vector $\vec{h}_2$, we compute

$$\begin{aligned} id\,(x+1) &= x+1 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 1\vec{e}_0 + 1\vec{e}_1 + 0\vec{e}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{H\to E} = \begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix}$$

Bases $E \to H$

Given the previous observation

$$\mathcal{R}_{G\to E}\mathcal{R}_{E\to G} = I$$

to compute

$$\mathcal{R}_{E\to H}$$

we simply need to find the inverse of matrix

$$\mathcal{R}_{H\to E} = \begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix}$$

which is

$$\mathcal{R}_{E\to H} = \begin{pmatrix} -1 & 1 & 1 \\ 3 & -3 & -2 \\ -4 & 5 & 2 \end{pmatrix}$$

Alternatively we can go the "long way": for basis vector $\vec{e}_0$, we compute

$$\begin{aligned} id\,(1) &= 1 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= -1\vec{e}_0 + 3\vec{e}_1 - 4\vec{e}_2 = \begin{pmatrix} -1 \\ 3 \\ -4 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

For basis vector $\vec{e}_1$, we compute

$$\begin{aligned} id\,(x) &= x \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 1\vec{e}_0 - 3\vec{e}_1 + 5\vec{e}_2 = \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

For basis vector $\vec{e}_2$, we compute

$$\begin{aligned} id\left(x^2\right) &= x^2 \\ &= x_0\vec{e}_0 + x_1\vec{e}_1 + x_2\vec{e}_2 \\ &= 1\vec{e}_0 - 2\vec{e}_1 + 2\vec{e}_2 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}_E \end{aligned}$$

where the coordinates are obtained by solving

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Solution is

$$\mathcal{R}_{E\to H} = \begin{pmatrix} -1 & 1 & 1 \\ 3 & -3 & -2 \\ -4 & 5 & 2 \end{pmatrix}$$

Bases $H \to G$

At this stage we will not go the long way. We already know $\mathcal{R}_{H\to E}$ and $\mathcal{R}_{E\to G}$. Naturally

$$\begin{aligned} \mathcal{R}_{H\to G} &= \mathcal{R}_{E\to G}\mathcal{R}_{H\to E} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 1 \\ -2 & -1 & 0 \\ 1 & -1 & -1 \end{pmatrix} \end{aligned}$$

Bases $G \to H$

As in the previous section we either compute

$$\mathcal{R}_{G\to H} = \mathcal{R}_{E\to H}\mathcal{R}_{G\to E}$$

which we already have computed or we fine the inverse of

$$\mathcal{R}_{H\to G}$$

in either case we get

$$\mathcal{R}_{G\to H} = \begin{pmatrix} 1 & 2 & 1 \\ -2 & -5 & -2 \\ 3 & 7 & 2 \end{pmatrix}$$

3.8.3 Invertible matrices

In the above examples if we want the matrix of the identity transformation $\mathcal{R}_{C\to R}$ we solve a system of linear equation $A\vec{x} = b$. If we want to get $\mathcal{R}_{R\to C}$ then we solve a system of linear equations that is $B\vec{x} = c$ and their relation is that $AB = I$. For any invertible matrix A we can do that. And anytime we have a change of basis matrix its columns (and rows) must be linearly independent so the matrix is invertible. This is the underlying statement of the following

Theorem 50. *A matrix changes basis if and only if it is non-singular.*

3.8.4 Matrix multiplication

The above Theorem 50 is for change of basis but if we combine change of basis with matrix representation of linear transformations we get the following picture:

$$\begin{array}{ccccc} \mathbf{V}_B & \to & \phi_{B\to D} & \to & \mathbf{W}_D \\ \downarrow & & & & \downarrow \\ id_{B\to E} & & & & id_{D\to C} \\ \downarrow & & & & \downarrow \\ \mathbf{V}_E & \to & \phi_{E\to C} & \to & \mathbf{W}_C \end{array}$$

For example for the derivative map from Section 3.7.1 we know

$$\mathcal{R}_{B\to E}, \mathcal{R}_{B\to G}, \mathcal{R}_{A\to E}, \mathcal{R}_{A\to H}$$

from Section 3.8.1 we know

$$\mathcal{R}_{A\to B}, \mathcal{R}_{B\to A}$$

from Section 3.8.2 we know

$$\mathcal{R}_{E\to G}, \mathcal{R}_{G\to E}, \mathcal{R}_{E\to H}, \mathcal{R}_{H\to E}, \mathcal{R}_{H\to G}, \mathcal{R}_{G\to H}$$

Suppose we did not know $\mathcal{R}_{A\to H}$ but we knew

- $\mathcal{R}_{B\to E}$ (very easy)
- $\mathcal{R}_{A\to B}$ (very easy)

- $\mathcal{R}_{E\to H}$ some computations

then to compute $\mathcal{R}_{A\to H}$ we need to use

$$\mathcal{R}_{A\to H} = \mathcal{R}_{E\to H}\mathcal{R}_{B\to E}\mathcal{R}_{A\to B}$$

which is just multiplication of matrices. So what is $\mathcal{R}_{A\to G}$? We use

$$\begin{aligned}\mathcal{R}_{A\to G} &= \mathcal{R}_{E\to G}\mathcal{R}_{B\to E}\mathcal{R}_{A\to B} \\ &= \begin{pmatrix} 2 & 0 & 3 & 1 \\ 2 & 2 & 3 & 1 \\ -7 & 1 & -3 & 1 \end{pmatrix}\end{aligned}$$

Chapter 4

Determinant

4.1 Definitions and properties

Definition 52. *A* $n \times n$-determinant *is a function*

$$\det : \mathcal{M}_{n\times n} \to \mathbb{K}$$

such that

$$\det(EA) = \det(E)\det(A)$$

for an elementary row operation matrix E *and any matrix* A*, with* $E, A \in \mathcal{M}_{n\times n}$*. Furthermore*

1. $\det(E) = 1$ *if* E *is linear combination;*
2. $\det(E) = -1$ *if* E *is a swap;*
3. $\det(E) = k$ *if* E *is rescaling;*
4. $\det(I) = 1$ *for the identity matrix of order* n*.*

We often write $|A|$ for $\det(A)$. Let the vectors $\vec{\rho}$ represent the rows of the matrix. The conditions then are written as:

$$\textbf{Row combination}\quad \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_i \\ \vec{\rho}_j + k\vec{\rho}_j \\ \vec{\rho}_i \\ \vdots \\ \vec{\rho}_n \end{pmatrix} \text{ for } i \neq j$$

$$\det\left[\begin{pmatrix} 1 & 0\dots & & & & & & \dots 0 & 0 \\ & & & \ddots & & & & & \\ 0 & \dots 0 & 0 & 0\dots & 1 & 0 & 0 & \dots 0 & 0 \\ 0 & \dots 0 & k & 0\dots & 0 & 1 & 0 & \dots 0 & 0 \\ 0 & \dots 0 & 0 & 0\dots & 0 & 0 & 1 & \dots 0 & 0 \\ & & & & & & & \ddots & \\ 0 & 0\dots & & & & & & \dots 0 & 1 \end{pmatrix}\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix}\right] = \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_i \\ \vec{\rho}_j + k\vec{\rho}_j \\ \vec{\rho}_i \\ \vdots \\ \vec{\rho}_n \end{pmatrix}$$

For example:

$$\begin{aligned} \det\begin{pmatrix} 5 & -8 & 1 \\ 0 & -\frac{1}{5} & \frac{2}{5} \\ -4 & 7 & -1 \end{pmatrix} &= \det\left(\begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{5} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix}\right) \\ &= \underbrace{\det\begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{5} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=1}\det\begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix} \\ &= 1\det\begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix} \end{aligned}$$

Swap $\det\begin{pmatrix}\vec\rho_1\\ \vdots\\ \vec\rho_i\\ \vdots\\ \vec\rho_j\\ \vdots\\ \vec\rho_n\end{pmatrix} = -\det\begin{pmatrix}\vec\rho_1\\ \vdots\\ \vec\rho_j\\ \vdots\\ \vec\rho_i\\ \vdots\\ \vec\rho_n\end{pmatrix}$ for $i \neq j$

$$\det\left[\begin{pmatrix}
1 & 0\ldots & & & & & & & & & & & \ldots 0 & 0\\
 & & \ddots & & & & & & & & & & & \\
0 & \ldots 0 & 0 & 0\ldots & 1 & 0 & 0 & & & & & & \ldots 0 & 0\\
0 & \ldots 0 & 1 & 0\ldots & 0 & 0 & 0 & & & & & & \ldots 0 & 0\\
0 & \ldots 0 & 0 & 0\ldots & 0 & 0 & 1 & & & & & & \ldots 0 & 0\\
 & & & & & & & \ddots & & & & & & \\
0 & 0\ldots & & & & & & 1 & 0 & 0 & \ldots 0 & 0 & 0\ldots & \\
0 & 0\ldots & & & & & & 0 & 0 & 0 & \ldots 0 & 1 & 0\ldots & \\
0 & 0\ldots & & & & & & 0 & 0 & 1 & \ldots 0 & 0 & 0\ldots & \\
 & & & & & & & & & & \ddots & & & \\
0 & 0\ldots & & & & & & & & & & & \ldots 0 & 1
\end{pmatrix}\begin{pmatrix}\vec\rho_1\\ \vdots\\ \vec\rho_{i-1}\\ \vec\rho_i\\ \vec\rho_{i+1}\\ \vdots\\ \vec\rho_{j-1}\\ \vec\rho_j\\ \vec\rho_{j+1}\\ \vdots\\ \vec\rho_n\end{pmatrix}\right] = -\det\begin{pmatrix}\vec\rho_1\\ \vdots\\ \vec\rho_{i-1}\\ \vec\rho_j\\ \vec\rho_{i+1}\\ \vdots\\ \vec\rho_{j-1}\\ \vec\rho_i\\ \vec\rho_{j+1}\\ \vdots\\ \vec\rho_n\end{pmatrix}$$

For example:

$$\begin{aligned}
\det\begin{pmatrix}3 & -5 & 1\\ 5 & -8 & 1\\ -4 & 7 & -1\end{pmatrix} &= \det\left(\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}5 & -8 & 1\\ 3 & -5 & 1\\ -4 & 7 & -1\end{pmatrix}\right)\\
&= \underbrace{\det\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}}_{=-1}\det\begin{pmatrix}5 & -8 & 1\\ 3 & -5 & 1\\ -4 & 7 & -1\end{pmatrix}\\
&= -1\det\begin{pmatrix}5 & -8 & 1\\ 3 & -5 & 1\\ -4 & 7 & -1\end{pmatrix}
\end{aligned}$$

Rescaling $\det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ k\vec{\rho}_i \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = k \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_i \\ \vdots \\ \vec{\rho}_n \end{pmatrix}$ for any scalar k

$$\det \left[\begin{pmatrix} 1 & 0\ldots & & & & \ldots 0 & 0 \\ & \ddots & & \vdots & & & \\ 0 & 0\ldots & 1 & 0 & 0 & \ldots 0 & 0 \\ 0 & 0\ldots & 0 & k & 0 & \ldots 0 & 0 \\ 0 & 0\ldots & 0 & 0 & 1 & \ldots 0 & 0 \\ & & & \vdots & & \ddots & \\ 0 & 0\ldots & & & & \ldots 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} \right] = \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ k\vec{\rho}_i \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix}$$

For example:

$$\begin{aligned} \det \begin{pmatrix} 5 & -8 & 1 \\ -15 & 25 & -5 \\ -4 & 7 & -1 \end{pmatrix} &= \det \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix} \right) \\ &= \underbrace{\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=-5} \det \begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix} \\ &= -5 \det \begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix} \end{aligned}$$

Theorem 51. *Condition 1 and Condition 3 imply Condition 2.*

Proof.

$$\begin{vmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_{j-1} \\ \vec{\rho}_j \\ \vec{\rho}_{j+1} \\ \vdots \\ \vec{\rho}_n \end{vmatrix} = \begin{vmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i + \vec{\rho}_j \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_{j-1} \\ \vec{\rho}_j \\ \vec{\rho}_{j+1} \\ \vdots \\ \vec{\rho}_n \end{vmatrix} = \begin{vmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i + \vec{\rho}_j \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_{j-1} \\ \vec{\rho}_j - \vec{\rho}_i - \vec{\rho}_j \\ \vec{\rho}_{j+1} \\ \vdots \\ \vec{\rho}_n \end{vmatrix} = \begin{vmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i + \vec{\rho}_j \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_{j-1} \\ -\vec{\rho}_i \\ \vec{\rho}_{j+1} \\ \vdots \\ \vec{\rho}_n \end{vmatrix} = \begin{vmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_i + \vec{\rho}_j - \vec{\rho}_i \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_{j-1} \\ -\vec{\rho}_i \\ \vec{\rho}_{j+1} \\ \vdots \\ \vec{\rho}_n \end{vmatrix} = \begin{vmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_j \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_{j-1} \\ -\vec{\rho}_i \\ \vec{\rho}_{j+1} \\ \vdots \\ \vec{\rho}_n \end{vmatrix} = - \begin{vmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_j \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_{j-1} \\ \vec{\rho}_i \\ \vec{\rho}_{j+1} \\ \vdots \\ \vec{\rho}_n \end{vmatrix}$$

$\square$

Swap (as a matrix) equals the product of linear combinations and rescaling. For example:

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

So if we apply the above matrix equality to determinants we get

$$\begin{aligned}
\det \begin{pmatrix} -2 & -6 & -6 & 3 \\ 0 & -5 & 0 & -2 \\ 0 & -1 & -1 & 0 \\ -2 & 0 & -5 & 1 \end{pmatrix} &= \underbrace{\det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}_{=-1} \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ 0 & -5 & 0 & -2 \\ 0 & -1 & -1 & 0 \\ 2 & 0 & 5 & -1 \end{pmatrix} \\
= (-1) \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ 0 & -5 & 0 & -2 \\ 0 & -1 & -1 & 0 \\ 2 & 0 & 5 & -1 \end{pmatrix} &= (-1) \underbrace{\det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{=1} \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ -2 & -5 & -5 & -1 \\ 0 & -1 & -1 & 0 \\ 2 & 0 & 5 & -1 \end{pmatrix} \\
= (-1) \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ -2 & -5 & -5 & -1 \\ 0 & -1 & -1 & 0 \\ 2 & 0 & 5 & -1 \end{pmatrix} &= (-1) \underbrace{\det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}}_{=1} \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ -2 & -5 & -5 & -1 \\ 0 & -1 & -1 & 0 \\ 0 & -5 & 0 & -2 \end{pmatrix} \\
= (-1) \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ -2 & -5 & -5 & -1 \\ 0 & -1 & -1 & 0 \\ 2 & 0 & 5 & -1 \end{pmatrix} &= (-1) \underbrace{\det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{=1} \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ -2 & 0 & -5 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -5 & 0 & -2 \end{pmatrix} \\
&= (-1) \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ -2 & 0 & -5 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -5 & 0 & -2 \end{pmatrix} \\
&= \det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \det \begin{pmatrix} -2 & -6 & -6 & 3 \\ -2 & 0 & -5 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -5 & 0 & -2 \end{pmatrix}
\end{aligned}$$

Theorem 52. *If a matrix A has a row of zeros then* $\det(A) = 0$.

Proof. Use scalar multiplication property with $k = 0$. □

Theorem 53. $\det(A) = 0$ *if and only if* $\vec{\rho_1}, \ldots, \vec{\rho_n}$ *are linearly dependent.*

Proof. Suppose that $\vec{\rho}_1, \ldots, \vec{\rho}_n$ are linearly dependent. Then there exist coefficients not all zero such that

$$\alpha_1\vec{\rho}_1 + \cdots + \alpha_n\vec{\rho}_n = \vec{0}$$

let $\alpha_k \neq 0$. Then

$$\alpha_k \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_k \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \alpha_k\vec{\rho}_k \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \alpha_k\vec{\rho}_k + \sum_{i=1, i\neq k}^{n} \alpha_i\vec{\rho}_i \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{0} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = 0$$

Since $\alpha_k \neq 0$ it follows that the determinant is zero.

Assume now the rows are linearly independent (and so are the columns). By Theorem 6 using the matrix representation of Gaussian operations we can write $A^{-1} = E_m \ldots E_1$ where each E_i is an elemetary matrix and all scaling operations do not involve a scaling by zero. Then

$$1 = \det(I) = \det(A^{-1}A) = \det(E_m \ldots E_1 A) = \det(E_m) \ldots \det(E_1)\det(A)$$

since the right hand side is non-zero the left hand side is also non zero; and therefore $\det(A) \neq 0$.

$\square$

On the one hande if we have an invertible matrix (i.e., rows are linearly independent) the matrix is a product of elementary matrices with no scaling by zero for example:

$$\begin{pmatrix} 0 & 1 & -3 \\ 0 & 0 & 1 \\ 1 & -2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix} \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \hookleftarrow$$

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

If we apply the determinant definition on the right hand side each determinant is either one, negative one or a scaling factor k that is different from zero. Thus we have a product of non-zero values and the result is non-zero determinant.

On the other hand if we have linearly dependent rows, by performing linear combinations we can get a row of zeros thus the determinant is zero for example:

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

On the right hand side there a matrix with rows of zeroes so its determinant is zero so the determinant of the matrix on the left side must be zero.

Theorem 54. $\det(A)$ *is unique.*

Proof. If columns of A are linear dependent then $\det(A) = 0 = \tilde{\det}(A)$. If colums of A are linearly independent then from

$$\begin{aligned}
\tilde{\det}(A) &= \tilde{\det}(E_m \dots E_1) \\
&= \tilde{\det}(E_m) \dots \tilde{\det}(E_1) \\
&= \det(E_m) \dots \det(E_1) \\
&= \det(E_m \dots E_1) \\
&= \det(A)
\end{aligned}$$

□

Theorem 55. $\det(AB) = \det(A)\det(B)$

Proof. If $\det(B) = 0$ then its rows (and by the rank) its rows are linearly dependent. In $C = AB$ the rows of C are linear combinations of the rows of B. Then the span of the rows of C is a subset of the span of the rows of B and therefore the number of linearly independent rows of C cannot exceed the number of linearly independent rows of B. Thus the rows of C cannot be linearly independent. Thus $\det(C) = 0$ and the theorem holds in this case. Similarly if $\det(A) = 0$ then the columns of A are linearly dependent and by a similar argument $\det(C) = 0$. Assume now that $\det(A) \neq 0$ and $\det(B) \neq 0$. Then as in the above theorem we have

$$\det(A) = \det(E_m) \dots \det(E_1)$$

$$\det(B) = \det(E'_k) \dots \det(E'_1)$$

and

$$\det(AB) = \det(E_m \dots E_1 E'_k \dots E'_1) = \det(E_m) \dots \det(E_1)\det(E'_k) \dots \det(E'_1) = \det(A)\det(B)$$

Note that in the above we do *not* use the fact that $\det(AB) = \det(AB)$ we simply use the definition where $\det(EA) = \det(E)\det(A)$ for any elementary matrix E. □

Recall that by Theorem 6 we have that

> Suppose the rows (columns) of square A are linearly independent then A can be written as a product of elementary matrices.

Theorem 56. $\det(A) = \det(A^T)$

Proof. If A is an elementary matrix then the result follows as:

1. Rescaling: if E is a rescaling matrix then $E^T = E$ and therefore $\det(E) = \det(E^T)$;
2. Swap: if E is a swap matrix then $E^T = E$ and therefore $\det(E) = \det(E^T)$;

3. row combination: if E is a matrix where k times row i is added to row j then E^T is the matrix k times row j is added to row i and therefore $\det(E^T) = 1$ where and therefore $\det(E) = \det(E^T)$; with details if

$$E = \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\rho}_j + k\vec{\rho}_j \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} \qquad \text{then} \qquad E^T = \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_j \\ \vec{\rho}_j + k\vec{\rho}_i \\ \vec{\rho}_j \\ \vdots \\ \vec{\rho}_n \end{pmatrix}$$

for $i \neq j$. Alternative write-up: if

$$E = \left(\begin{array}{cccccccccccc}
1 & 0 & \dots & & & & & & & & \dots 0 & 0 \\
0 & 1 & & & & & & & & & & \\
 & & \ddots & & & & & & & & & \\
0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \dots 0 & 0 \\
\vdots & & & & & & \ddots & & & & & \vdots \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & k & 0 & \dots & 0 & 1 & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 1 & \dots 0 & 0 \\
 & & & & & & & & & & \ddots & \\
0 & 0 & \dots & & & & & & & & \dots 0 & 1
\end{array}\right)$$

then

$$E^T = \left(\begin{array}{cccccccccccc}
1 & 0 & \dots & & & & & & & & \dots 0 & 0 \\
0 & 1 & & & & & & & & & & \\
 & & \ddots & & & & & & & & & \\
0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & k & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \dots 0 & 0 \\
\vdots & & & & & & \ddots & & & & & \vdots \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots 0 & 0 \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 1 & \dots 0 & 0 \\
 & & & & & & & & & & \ddots & \\
0 & 0 & \dots & & & & & & & & \dots 0 & 1
\end{array}\right)$$

for $i \neq j$.

Now if A has linearly dependent rows it has linearly dependend columns by Theorem 21 its rows are also linearly dependent. Then the columns of A^T are linearly dependent and $\det(A^T) = 0 = \det(A)$. Suppose now A has linearly independent rows and by Theorem 6 A can be represented as product of elementary matrices. Say

$$A = E_1 E_2 \dots E_s$$

then

$$A^T = E_s^T E_{s-1}^T \dots E_2^T E_1^T.$$

Hence by Theorem 55 we have

$$\begin{aligned}
\det(A^T) &= \det(E_s^T)\det(E_{s-1}^T)\dots\det(E_2^T)\det(E_1^T) \\
&= \det(E_s)\det(E_{s-1})\dots\det(E_2)\det(E_1) \\
&= \det(E_1)\det(E_2)\dots\det(E_{s-1})\det(E_s) \\
&= \det(E_1 E_2 \dots E_{s-1} E_s) = \det(A)
\end{aligned}$$

Which concludes the argument. □

4.2 Towards existence

In the previous section we established that if there is determinant function then this function is unique. But does it exists? Consider the (invertible) matrix

$$\begin{pmatrix} 0 & 1 & -3 \\ 0 & 0 & 1 \\ 1 & -2 & 5 \end{pmatrix}$$

it can be written as a product of elementary matrices as:

$$\begin{pmatrix} 0 & 1 & -3 \\ 0 & 0 & 1 \\ 1 & -2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix}\begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \hookleftarrow$$

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

as well as

$$\begin{pmatrix} 0 & 1 & -3 \\ 0 & 0 & 1 \\ 1 & -2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \hookleftarrow$$

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

Existence question relates to how do we know that from $A = E_m \dots E_1$ and $A = E'_k \dots E'_1$ that

$$\det(E_m)\dots\det(E_1) = \det(E'_k)\dots\det(E'_1).$$

May be the product on the left and the right are not equal. Another example for the matrix

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

how many swaps do we need to get the identity matrix? May be we can do it with 8 swaps and at the same time we can do it with 5 swaps. In one case we have determinant positive one and in the other negative one.

Theorem 57. $\det\begin{pmatrix} \vec\rho_1 \\ \vdots \\ \vec\rho_{i-1} \\ \vec u + \vec v \\ \vec\rho_{i+1} \\ \vdots \\ \vec\rho_n \end{pmatrix} = \det\begin{pmatrix} \vec\rho_1 \\ \vdots \\ \vec\rho_{i-1} \\ \vec u \\ \vec\rho_{i+1} \\ \vdots \\ \vec\rho_n \end{pmatrix} + \det\begin{pmatrix} \vec\rho_1 \\ \vdots \\ \vec\rho_{i-1} \\ \vec v \\ \vec\rho_{i+1} \\ \vdots \\ \vec\rho_n \end{pmatrix}$

Proof. If $\vec\rho_1, \dots, \vec\rho_{i-1}, \vec\rho_{i+1}, \dots, \vec\rho_n$ are linearly dependent then all determinants are zero and the result follows. Suppose now they are linearly independent. Then we can find a vector $\vec\beta$ such that $\vec\rho_1, \dots, \vec\rho_{i-1}, \vec\beta, \vec\rho_{i+1}, \dots, \vec\rho_n$ are linearly independent. Since there are n of them they span all of $\mathbb{K}^n$ and therefore

$$\begin{aligned} \vec u &= u_1\vec\rho_1 + \dots + u_{i-1}\vec\rho_{i-1} + u_i\vec\beta + u_{i+1}\vec\rho_{i+1} + \dots + u_n\vec\rho_n \\ \vec v &= v_1\vec\rho_1 + \dots + v_{i-1}\vec\rho_{i-1} + v_i\vec\beta + v_{i+1}\vec\rho_{i+1} + \dots + v_n\vec\rho_n \\ \vec u + \vec v &= (u_1+v_1)\vec\rho_1 + \dots + (u_{i-1}+v_{i-1})\vec\rho_{i-1} \\ &\quad +(u_i+v_i)\vec\beta + (u_{i+1}+v_{i+1})\vec\rho_{i+1} + \dots + (u_n+v_n)\vec\rho_n \end{aligned}$$

Then for $j = 1 \dots i-1$ and $j = i+1, \dots n$ we have

$$\det\begin{pmatrix} \vec\rho_1 \\ \vdots \\ \vec\rho_{i-1} \\ \vec u + \vec v = \vec w_0 \\ \vec\rho_{i+1} \\ \vdots \\ \vec\rho_n \end{pmatrix} = \det\begin{pmatrix} \vec\rho_1 \\ \vdots \\ \vec\rho_{i-1} \\ \vec w_{j-1} - (u_j+v_j)\vec\rho_j \\ \vec\rho_{i+1} \\ \vdots \\ \vec\rho_n \end{pmatrix}$$

At the end we obtain

$$
\begin{aligned}
\det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{u}+\vec{v} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} &= \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ (u_i+v_i)\vec{\beta} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = (u_i+v_i)\det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\beta} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} \\
&= u_i\det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\beta} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} + v_i\det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\beta} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix}
\end{aligned}
$$

Consider

$$
\begin{aligned}
u_i\det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\beta} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} &= \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ u_i\vec{\beta} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ u_i\vec{\beta}+u_1\vec{\rho}_1 \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} \\
&= \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ u_i\vec{\beta}+u_1\vec{\rho}_1+u_2\vec{\rho}_2 \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = \cdots = \det\begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{u} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix}
\end{aligned}
$$

Similarly,

$$v_i \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{\beta} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix} = \cdots = \det \begin{pmatrix} \vec{\rho}_1 \\ \vdots \\ \vec{\rho}_{i-1} \\ \vec{v} \\ \vec{\rho}_{i+1} \\ \vdots \\ \vec{\rho}_n \end{pmatrix}$$

And substituting back we get the desired result. $\square$

Permutation. A permutation of n is a bijective function with domain and range the set of numbers $1, \ldots, n$,

$$\phi : [1, n] \to [1, n].$$

For example

n	1	2	3	4
$\phi(n)$	2	4	3	1

is a permutation of 4. The permutation is given in table form e.g., $\phi(2) = 4$. The matrix form of a permutation of n is an $n \times n$ matrix where in row i all elements are zero except the entry in column $\phi(i)$ which is one. For the permutation of 4 given above the matrix is

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

Example: From

$$\begin{pmatrix} 5 & -8 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -8 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$

we have the expansion

$$\begin{aligned}
\det\begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix} &= \det\begin{pmatrix} 5 & 0 & 0 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix} + \det\begin{pmatrix} 0 & -8 & 0 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix} + \det\begin{pmatrix} 0 & 0 & 1 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix} \\
&= \det\begin{pmatrix} 5 & 0 & 0 \\ 3 & 0 & 0 \\ -4 & -7 & -1 \end{pmatrix} + \det\begin{pmatrix} 5 & 0 & 0 \\ 0 & -5 & 0 \\ -4 & -7 & -1 \end{pmatrix} + \det\begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 1 \\ -4 & -7 & -1 \end{pmatrix} \\
&\quad + \det\begin{pmatrix} 0 & -8 & 0 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix} + \det\begin{pmatrix} 0 & 0 & 1 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix} \\
&= \det\begin{pmatrix} 5 & 0 & 0 \\ 3 & 0 & 0 \\ -4 & 0 & 0 \end{pmatrix} + \det\begin{pmatrix} 5 & 0 & 0 \\ 3 & 0 & 0 \\ 0 & -7 & 0 \end{pmatrix} + \det\begin{pmatrix} 5 & 0 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
&\quad + \det\begin{pmatrix} 5 & 0 & 0 \\ 0 & -5 & 0 \\ -4 & -7 & -1 \end{pmatrix} + \det\begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 1 \\ -4 & -7 & -1 \end{pmatrix} \\
&\quad + \det\begin{pmatrix} 0 & -8 & 0 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix} + \det\begin{pmatrix} 0 & 0 & 1 \\ 3 & -5 & 1 \\ -4 & -7 & -1 \end{pmatrix}
\end{aligned}$$

When we take the same element from the first row we get a matrix where the first two rows are linearly dependent hence the determinant is zero. So the only contribution comes from determinants where we take elements from different columns (and thus keep linearly independence). So we end up with the formula for the determinant

$$\begin{aligned}
\det(A) &= \sum_{\text{permutations } \phi} a_{1\phi(1)} \dots a_{n\phi(n)} \det(P_\phi) && (4.1) \\
&= \sum_{i=1}^{n} (-1)^{i+c} a_{ic} \det\left(A(i|c)\right) && (4.2) \\
&= \sum_{i=1}^{n} (-1)^{r+i} a_{ri} \det\left(A(r|i)\right) && (4.3)
\end{aligned}$$

Here $A(i|j)$ is a submatrix obtained from A by removing row i and column j. The value $\det(A(i|j)$ is called the minor of a_{ij}, and the value $(-1)^{k+i} \det\left(A(i|k)\right)$ is called the cofactor of a_{ij}. Equation 4.2 (column expansion) and Equation 4.3 (row expansion) can be obtained from Equation 4.1 via tedious manipulations.

Remark. While determinant function has domain the set of square $n \times n$ matrices, when talking about $A(i|j)$ in general A need not be a square matrix. The notation $A(i|j)$ is often used for any $m \times n$ matrix. Cofactors are defined via a determinant therefore cofactors are defined only for square matrices.

Example: Column expansion by first column

$$
\begin{aligned}
\det \begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix} &= (5)\,(-1)^{1+1} \det \begin{pmatrix} -5 & 1 \\ 7 & -1 \end{pmatrix} \\
&\quad + (3)\,(-1)^{2+1} \det \begin{pmatrix} -8 & 1 \\ 7 & -1 \end{pmatrix} \\
&\quad + (-4)\,(-1)^{3+1} \det \begin{pmatrix} -8 & 1 \\ -5 & 1 \end{pmatrix} \\
&= (5)\,(-1)^{1+1}\,(-2) \\
&\quad + (3)\,(-1)^{2+1}\,(1) \\
&\quad + (-4)\,(-1)^{3+1}\,(-3) \\
&= -1
\end{aligned}
$$

Observe that repeatedly applying columns expansion by first column implies that the determinant of a triangular matrix is the product of diagonal entries.

Example: Row expansion by second row

$$
\begin{aligned}
\det \begin{pmatrix} 5 & -8 & 1 \\ 3 & -5 & 1 \\ -4 & 7 & -1 \end{pmatrix} &= (3)\,(-1)^{2+1} \det \begin{pmatrix} -8 & 1 \\ 7 & -1 \end{pmatrix} \\
&\quad + (-5)\,(-1)^{2+2} \det \begin{pmatrix} 5 & 1 \\ -4 & -1 \end{pmatrix} \\
&\quad + (1)\,(-1)^{2+3} \det \begin{pmatrix} 5 & -8 \\ -4 & 7 \end{pmatrix} \\
&= (3)\,(-1)^{2+1}\,(1) \\
&\quad + (-5)\,(-1)^{2+2}\,(-1) \\
&\quad + (1)\,(-1)^{2+3}\,(3) \\
&= -1
\end{aligned}
$$

4.3 Determinant of a permutation

From Equation 4.1 to show that determinant is well defined function we need to show it is well defined for permutations. These are matrices that have exactly one entry one in each column and row and all other entries are zero.

Definition 53. $\phi = (\phi(1), \phi(2), \ldots, \phi(n))$. *In a permutation matrix* $P_\phi = \begin{pmatrix} \vdots \\ \rho_{\phi(k)} \\ \vdots \\ \rho_{\phi(l)} \\ \vdots \end{pmatrix}$ *two rows k and l with $k < l$ are an inversion if and only if $\phi(k) > \phi(l)$.*

Example: the permutation

n	1	2	3	4
$\phi(n)$	2	4	3	1

has four inversions: $\phi(1) > \phi(4)$, $\phi(2) > \phi(3)$, $\phi(2) > \phi(4)$ and $\phi(3) > \phi(4)$.

Theorem 58. *A row swap changes in a permutation matrix changes the total number of inversions either from even to odd or from odd to even.*

Proof. Suppose we swap two rows that are adjacent

$$\begin{pmatrix} \vdots \\ \rho_{\phi(s)} \\ \vdots \\ \rho_{\phi(k)} \\ \rho_{\phi(k+1)} \\ \vdots \\ \rho_{\phi(t)} \\ \vdots \end{pmatrix} \leftrightarrow \begin{pmatrix} \vdots \\ \rho_{\phi(s)} \\ \vdots \\ \rho_{\phi(k+1)} \\ \rho_{\phi(k)} \\ \vdots \\ \rho_{\phi(t)} \\ \vdots \end{pmatrix}$$

Then $\phi(s)\phi(k)$ is an inversion in the first matrix if and only if it is an inversion in the second matrix. Similarly $\phi(s)\phi(k+1)$ is an inversion in the first matrix if and only if it is an inversion in the second matrix; $\phi(s+1)\phi(k)$ is an inversion in the first matrix if and only if it is an inversion in the second matrix; $\phi(s+1)\phi(k+1)$ is an inversion in the first matrix if and only if it is an inversion in the second matrix. However, $\phi(k)\phi(k+1)$ is an inversion of the first matrix if and only if $\phi(k+1), \phi(k)$ is *not* and inversion in the second matrix. Thus a swap of two adjacent rows changes the parity of inversions. Consider now

$$\begin{pmatrix} \vdots \\ \rho_{\phi(s)} \\ \vdots \\ \rho_{\phi(k)} \\ \vdots \\ \rho_{\phi(l)} \\ \vdots \\ \rho_{\phi(t)} \\ \vdots \end{pmatrix} \leftrightarrow \begin{pmatrix} \vdots \\ \rho_{\phi(s)} \\ \vdots \\ \rho_{\phi(k)} \\ \rho_{\phi(l)} \\ \vdots \\ \rho_{\phi(l-1)} \\ \vdots \\ \rho_{\phi(t)} \\ \vdots \end{pmatrix} \leftrightarrow \begin{pmatrix} \vdots \\ \rho_{\phi(s)} \\ \vdots \\ \rho_{\phi(l)} \\ \rho_{\phi(k)} \\ \vdots \\ \rho_{\phi(l-1)} \\ \vdots \\ \rho_{\phi(t)} \\ \vdots \end{pmatrix} \leftrightarrow \begin{pmatrix} \vdots \\ \rho_{\phi(s)} \\ \vdots \\ \rho_{\phi(l)} \\ \vdots \\ \rho_{\phi(l-1)} \\ \rho_{\phi(k)} \\ \vdots \\ \rho_{\phi(t)} \\ \vdots \end{pmatrix} \leftrightarrow \begin{pmatrix} \vdots \\ \rho_{\phi(s)} \\ \vdots \\ \rho_{\phi(l)} \\ \vdots \\ \rho_{\phi(k)} \\ \vdots \\ \rho_{\phi(t)} \\ \vdots \end{pmatrix}$$

We perform the swaps by moving row l up to row k by performing swaps of adjacent rows. The number of swap that is required is $k - l$. Then moving row k, which is now at position $k + 1$ to row l, which requires a total of $k - 1 - l$ swaps. The total number of swaps is then $2k - 2l - 1$ which is an odd number. Thus swapping two rows changes the parity of number of inversions. $\square$

Theorem 59. *If a permutation has odd number of inversions then swapping to the identity matrix takes odd number of swaps. If a permutation has even number of swaps then swapping it to the identity matrix takes even number of swaps.*

Proof. Identity has zero number of swaps, hence to change odd number to zero requires odd swaps and changing even number to zero requires even number of swaps. $\square$

Definition 54. *The sign of a permutation is negative one if the number in inversions is odd, and positive one if the number of inversions is even.*

The following functions satisfies the determinant properties:

$$\begin{aligned} \det(A) &= \sum_{\text{permutations } \phi} a_{1\phi(1)} \dots a_{n\phi(n)} \det(P_\phi) \\ &= \sum_{\text{permutations } \phi} a_{1\phi(1)} \dots a_{n\phi(n)} sign(P_\phi) \end{aligned}$$

Indeed P_ϕ is an invertible matrix and as discussed above it can be written as product of swaps only. Then $P_\phi = E_1 E_2 \dots E_n$ and applying the determinant definition we obtain $\det(P_\phi) = \det(E_1)\det(E_2)\dots\det(E_n)$. Since the number n of swaps is always even or always odd for a fixed permutation matrix, the product on the right hand side always equals the sign of the permutation completing the argument.

Chapter 5

Eigenvalues and Eigenvectors

5.1 Motivation

Matrices are linear transformation. Under a given transformation some things remain unchanged. An important idea is to find what remains unchanged. Such information has extensive application from compression to page ranking to animations. We will next discuss the mathematics behind those ideas.

5.2 Eigenvectors

Definition 55 (eigenvalues and eigenvectors)**.** *Let A be a square matrix. A non-zero vector $\vec{u}$ is an* eigenvector *for A if $A\vec{u} = \lambda\vec{u}$ for some λ. The value λ is called* eigenvalue *for the eigenvector $\vec{u}$.*

We also use the term linear transformation instead of a square matrices. In that case we mean the matrix of the linear transformation from a basis B to the same basis B.

5.2.1 Examples

Consider

$$\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} -10 \\ -4 \end{pmatrix} = \begin{pmatrix} -20 \\ -8 \end{pmatrix} = 2 \begin{pmatrix} -10 \\ -4 \end{pmatrix}$$

Then $\begin{pmatrix} -10 \\ -4 \end{pmatrix}$ is eigenvector for $\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix}$ with corresponding eigenvalue 2.

Consider

$$\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix} \begin{pmatrix} -6 \\ -2 \\ -8 \end{pmatrix} = \begin{pmatrix} -6 \\ -2 \\ -8 \end{pmatrix} = 1 \begin{pmatrix} -6 \\ -2 \\ -8 \end{pmatrix}$$

Then $\begin{pmatrix} -6 \\ -2 \\ -8 \end{pmatrix}$ is eigenvector for $\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix}$ with corresponding eigenvalue 1.

Consider

$$\begin{pmatrix} -1 & 0 & 3 & 1 & 5 \\ 0 & -1 & 6 & 2 & 10 \\ 3 & -1 & -4 & -1 & -6 \\ -4 & 3 & -3 & -2 & -7 \\ -1 & 0 & 3 & 1 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \\ -1 \end{pmatrix}$$

Then $\begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \\ -1 \end{pmatrix}$ is eigenvector for $\begin{pmatrix} -1 & 0 & 3 & 1 & 5 \\ 0 & -1 & 6 & 2 & 10 \\ 3 & -1 & -4 & -1 & -6 \\ -4 & 3 & -3 & -2 & -7 \\ -1 & 0 & 3 & 1 & 5 \end{pmatrix}$ with corresponding eigenvalue 0.

5.2.2 Remarks

Observe that the problem $A\vec{x} = \lambda\vec{x}$ is not linear. For the three by three matrix above we get

$$\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix} \vec{x} = \lambda\vec{x}$$

or written explicitly we get

$$\begin{aligned} -9x_1 + 14x_2 + 4x_3 &= \lambda x_1 \\ 3x_1 - 2x_3 &= \lambda x_2 \\ -18x_1 + 22x_2 + 9x_3 &= \lambda x_3 \end{aligned}$$

Both $\vec{x}$ and λ are not known.

Furthermore, if we set $\vec{x} = \vec{0}$ then $A\vec{x} = \lambda\vec{x}$ becomes $A\vec{0} = \lambda\vec{0}$ which is satisfied for all values λ. Such solution is not that interesting. Similar to linear dependence and independence we are interested in non-trivial solutions for $\vec{x}$. Therefore, the definition requires that $\vec{u}$ is a non-zero vector, but allows for λ to be zero. That is 0 can be an eigenvalue but $\vec{0}$ cannot be an eigenvector.

5.3 Existence

The first questions that we need to address is: given an $n \times n$ matrix is it true that such a matrix always will have an eigenvector (and a corresponding eigenvalue).

5.3.1 Polynomials and matrices

For a $n \times n$ matrix M let

$$\begin{aligned}
M^0 &= I_n \\
M^1 &= M \\
M^2 &= MM \\
M^3 &= MMM \\
&\vdots \\
M^k &= \underbrace{M \cdots M}_{k \text{ times}}
\end{aligned}$$

With the above polynomials can be evaluated at matrices. For a polynomial

$$\begin{aligned}
p(x) &= a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k \\
&= a_0 x^0 + a_1 x^1 + a_2 x^2 + \cdots + a_k x^k \\
&= (x - x_0)^{r_0} (x - x_1)^{r_1} \cdots (x - x_t)^{r_t}
\end{aligned}$$

where $\forall i, r_i \geq 1, r_i \in \mathbb{Z}$ and $r_0 + r_1 + \cdots + r_t = k$, define

$$\begin{aligned}
p(M) &= a_0 M^0 + a_1 M^1 + a_2 M^2 + \cdots + a_k M^k \\
&= (M - x_1 I_n)^{r_1} (M - x_2 I_n)^{r_2} \cdots (M - x_t I_n)^{r_t}
\end{aligned}$$

Examples

For the matrix M and the corresponding polynomial $p(x)$ verify the evaluations below:

$$\begin{aligned}
M &= \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} \\
p(x) &= x^2 - 4 \\
&= 1 \begin{pmatrix} 6 & -5 \\ 2 & -1 \end{pmatrix} + 0 \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} - 4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= (x - 2) \cdot (x + 2) \\
&= \underbrace{\begin{pmatrix} 2 & -5 \\ 2 & -5 \end{pmatrix}}_{M-(2)I_2} \underbrace{\begin{pmatrix} 6 & -5 \\ 2 & -1 \end{pmatrix}}_{M-(-2)I_2} \\
&= \begin{pmatrix} 2 & -5 \\ 2 & -5 \end{pmatrix}
\end{aligned}$$

The same polynomial evaluated at a different matrix:

$$
\begin{aligned}
M &= \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix} \\
p(x) &= x^2 - 4 \\
&= 1 \begin{pmatrix} 12 & 13 & 20 \\ 4 & 5 & 7 \\ 15 & 16 & 25 \end{pmatrix} + 0 \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix} - 4 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= (x-2) \cdot (x+2) \\
&= \underbrace{\begin{pmatrix} 0 & 2 & 3 \\ 1 & -2 & 1 \\ 2 & 3 & 2 \end{pmatrix}}_{M-(2)I_3} \underbrace{\begin{pmatrix} 4 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & 3 & 6 \end{pmatrix}}_{M-(-2)I_3} \\
&= \begin{pmatrix} 8 & 13 & 20 \\ 4 & 1 & 7 \\ 15 & 16 & 21 \end{pmatrix}
\end{aligned}
$$

The same matrix evaluated at a different polynomial:

$$
\begin{aligned}
M &= \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix} \\
p(x) &= x^3 - 13x + 12 \\
&= 1 \begin{pmatrix} 77 & 84 & 129 \\ 27 & 29 & 45 \\ 96 & 105 & 161 \end{pmatrix} + 0 \begin{pmatrix} 12 & 13 & 20 \\ 4 & 5 & 7 \\ 15 & 16 & 25 \end{pmatrix} - 13 \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix} + 12 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= (x-3) \cdot (x-1) \cdot (x+4) \\
&= \underbrace{\begin{pmatrix} -1 & 2 & 3 \\ 1 & -3 & 1 \\ 2 & 3 & 1 \end{pmatrix}}_{M-(3)I_3} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 2 & 3 & 3 \end{pmatrix}}_{M-(1)I_3} \underbrace{\begin{pmatrix} 6 & 2 & 3 \\ 1 & 4 & 1 \\ 2 & 3 & 8 \end{pmatrix}}_{M-(-4)I_3} \\
&= \begin{pmatrix} 63 & 58 & 90 \\ 14 & 41 & 32 \\ 70 & 66 & 121 \end{pmatrix}
\end{aligned}
$$

5.3.2 Matrices as maps

Let M be an $n \times n$ matrix and $\vec{u}$ be a non zero vector with n components. M can be viewed as a linear transformation. Consider

$$\begin{aligned}
S_0 &= \left\{M^0\vec{u}\right\} \\
S_1 &= \left\{M^0\vec{u}, M^1\vec{u}\right\} \\
S_2 &= \left\{M^0\vec{u}, M^1\vec{u}, M^2\vec{u}\right\} \\
&\vdots \\
S_t &= \left\{M^0\vec{u}, M^1\vec{u}, M^2\vec{u}, \ldots, M^t\vec{u}\right\} \\
&\vdots
\end{aligned}$$

There is a minimum index $k \geq 1$ such that S_k is linearly dependent and for all $j < k$ the set S_j is linearly independent. Indeed S_0 is just the vector $\vec{u}$ that is non-zero, therefore S_0 is linearly independent. On the other hand any set of $i \geq n$ vectors in the n-dimensional vector space are linearly dependent thus $k \leq n$.

Let S_k be first set in the sequence of sets of vectors $S_0, S_1 \ldots$ that is linearly dependent. Adopting the notation

$$\vec{u}_i = M^i\vec{u}$$

which recursively means

$$\vec{u}_i = M\vec{u}_{i-1}$$

the set S_k is

$$S_k = \left\{M^0\vec{u}, M^1\vec{u}, M^2\vec{u}, \ldots, M^{k-1}\vec{u}, M^k\vec{u}\right\} = \{\vec{u_0}, \vec{u_1}, \vec{u_2}, \ldots, \vec{u_{k-1}}, \vec{u_k}\}$$

Since S_t is linearly dependent there are coefficients (not all zero):

$$\alpha_0\vec{u_0} + \alpha_1\vec{u_1} + \alpha_2\vec{u_2} + \cdots + \alpha_{k-1}\vec{u_{k-1}} + \alpha_k\vec{u_k} = \vec{0}$$

Observe that if $\alpha_k = 0$ then the set S_{k-1} is linearly dependent contradicting minimality of k; thus division by α_k is allowed

$$\frac{\alpha_0}{\alpha_k}\vec{u_0} + \frac{\alpha_1}{\alpha_k}\vec{u_1} + \frac{\alpha_2}{\alpha_k}\vec{u_2} + \cdots + \frac{\alpha_{k-1}}{\alpha_k}\vec{u_{k-1}} + \vec{u_k} = \vec{0}$$

Let

$$a_i = \frac{\alpha_i}{\alpha_k}$$

then

$$a_0\vec{u_0} + a_1\vec{u_1} + a_2\vec{u_2} + \cdots + a_{k-1}\vec{u_{k-1}} + \vec{u_k} = \vec{0}$$

By substituting back the notation for $\vec{u_i}$ the above equation implies:

$$\begin{aligned}
\vec{0} &= a_0\vec{u_0} + a_1\vec{u_1} + a_2\vec{u_2} + \cdots + a_{k-1}\vec{u_{k-1}} + \vec{u_k} \\
&= a_0M^0\vec{u} + a_1M^1\vec{u} + a_2M^2\vec{u} + \cdots + a_{k-1}M^{k-1}\vec{u} + M^k\vec{u} \\
&= \left(a_0M^0 + a_1M^1 + a_2M^2 + \cdots + a_{k-1}M^{k-1} + M^k\right)\vec{u} \\
&= p(M)\vec{u} \\
&= (M - \lambda_k I_n)(M - \lambda_{k-1} I_n)\cdots(M - \lambda_2 I_n)(M - \lambda_1 I_n)\vec{u}
\end{aligned}$$

Matrix multiplication is associative operation so the order of multiplication does not change the outcome of the computation. Performing the multiplication from right to left we have:

$$\begin{aligned}
\vec{z_0} &= \vec{u} \neq \vec{0} \\
\vec{z_1} &= (M - \lambda_1 I_n)\vec{u} \\
\vec{z_2} &= (M - \lambda_2 I_n)(M - \lambda_1 I_n)\vec{u} \\
&\vdots \\
\vec{z_{i-1}} &= (M - \lambda_{i-1} I_n)\cdots(M - \lambda_2 I_n)(M - \lambda_1 I_n)\vec{u} \\
\vec{z_i} &= (M - \lambda_i I_n)\underbrace{(M - \lambda_{i-1} I_n)\cdots(M - \lambda_2 I_n)(M - \lambda_1 I_n)\vec{u}}_{=\vec{z_{i-1}}} \\
&\vdots \\
\vec{z_{k-1}} &= (M - \lambda_{k-1} I_n)\cdots(M - \lambda_2 I_n)(M - \lambda_1 I_n)\vec{u} \\
\vec{0} = \vec{z_k} &= (M - \lambda_k I_n)(M - \lambda_{k-1} I_n)\cdots(M - \lambda_2 I_n)(M - \lambda_1 I_n)\vec{u}
\end{aligned}$$

We have found an eigenvector with corresponding eigenvalue: if $\vec{z_i} = \vec{0}$ and $\vec{z_{i-1}} \neq \vec{0}$ then λ_i is an eigenvalue and $\vec{z_{i-1}}$ is a corresponding eigenvector. Which shows that every $n \times n$ matrix has at least one eigenvector and corresponding eigenvalue.

Examples

Applying the above procedure to

$$\begin{aligned}
M &= \begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix} \\
\vec{u} &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

The sets S_i are

$$\begin{aligned}
S_0 &= \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\} \\
S_1 &= \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} -19 \\ -7 \\ -10 \end{pmatrix} \right\} \\
S_2 &= \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} -19 \\ -7 \\ -10 \end{pmatrix} \begin{pmatrix} -19 \\ -13 \\ 2 \end{pmatrix} \right\} \\
S_3 &= \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} -19 \\ -7 \\ -10 \end{pmatrix} \begin{pmatrix} -19 \\ -13 \\ 2 \end{pmatrix} \begin{pmatrix} -139 \\ -61 \\ -52 \end{pmatrix} \right\}
\end{aligned}$$

S_0 is linear independent and S_3 definitely linearly dependent. But there may be another set with index smaller that 3 that is linearly dependent. Following the above procedure check if S_1 is linearly dependent that is solve:

$$s_0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + s_1 \begin{pmatrix} -19 \\ -7 \\ -10 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

the set of solutions contains only the trivial solution so it is linearly independent. Next, is S_2 linearly independent, that is solve:

$$s_0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + s_1 \begin{pmatrix} -19 \\ -7 \\ -10 \end{pmatrix} + s_2 \begin{pmatrix} -19 \\ -13 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

the set of solutions contains only the trivial solution so it is linearly independent. The set S_3 is linearly dependent so the index we need is 3. Furthermore, the equation

$$s_0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + s_1 \begin{pmatrix} -19 \\ -7 \\ -10 \end{pmatrix} + s_2 \begin{pmatrix} -19 \\ -13 \\ 2 \end{pmatrix} + s_3 \begin{pmatrix} -139 \\ -61 \\ -52 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has non-trivial solution $\begin{pmatrix} s_0 & s_1 & s_2 & s_3 \end{pmatrix} = \begin{pmatrix} 6 & -5 & -2 & 1 \end{pmatrix}$. For the corresponding polynomial

$$\begin{aligned}
p(x) &= x^3 - 2x^2 - 5x + 6 \\
&= (x-3) \cdot (x-1) \cdot (x+2)
\end{aligned}$$

the vectors $\vec{z}_i$ are

$$
\begin{aligned}
\vec{z}_0 &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\
\vec{z}_1 &= \begin{pmatrix} -17 \\ -7 \\ -8 \end{pmatrix} = (M - (-2)I) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\
\vec{z}_2 &= \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix} = (M - (1)I) \begin{pmatrix} -17 \\ -7 \\ -8 \end{pmatrix} \\
\vec{z}_3 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (M - (3)I) \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix}
\end{aligned}
$$

The equation

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (M - (3)I) \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix}$$

rearranged as

$$M \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix} = 3I \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix}$$

equivalently

$$\begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix} \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix} = 3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix}$$

simplified

$$\begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix} \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix} = 3 \begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix}$$

identifies $\begin{pmatrix} -40 \\ -20 \\ -10 \end{pmatrix}$ as eigenvector with eigenvalue 3 for matrix $\begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix}$

The eigenvector that is computed depends on the initial vector $\vec{u}$ and the order of which roots of the polynomial are listed. Here is another example for the same matrix but with different initial vector:

$$
\begin{aligned}
M &= \begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix} \\
\vec{u} &= \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}
\end{aligned}
$$

The sets S_i are:

$$\begin{aligned}
S_0 &= \left\{ \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \right\} \\
S_1 &= \left\{ \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} 11 \\ 5 \\ 4 \end{pmatrix} \right\} \\
S_2 &= \left\{ \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} 11 \\ 5 \\ 4 \end{pmatrix} \begin{pmatrix} 35 \\ 17 \\ 10 \end{pmatrix} \right\} \\
S_3 &= \left\{ \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} 11 \\ 5 \\ 4 \end{pmatrix} \begin{pmatrix} 35 \\ 17 \\ 10 \end{pmatrix} \begin{pmatrix} 107 \\ 53 \\ 28 \end{pmatrix} \right\}
\end{aligned}$$

S_0 is linear independent and S_3 definitely linearly dependent. But there may be another set with index smaller that 3 that is linearly dependent. Following the above procedure check if S_1 is linearly dependent that is solve:

$$s_0 \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + s_1 \begin{pmatrix} 11 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

the set of solutions contains only the trivial solution so it is linearly independent. Next, is S_2 linearly independent, that is solve:

$$s_0 \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + s_1 \begin{pmatrix} 11 \\ 5 \\ 4 \end{pmatrix} + s_2 \begin{pmatrix} 35 \\ 17 \\ 10 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The set S_2 is linearly dependent. For example the above equation has a non-trivial solution $\begin{pmatrix} s_0 & s_1 & s_2 \end{pmatrix} = \begin{pmatrix} 3 & -4 & 1 \end{pmatrix}$. Thus the index k is 2. For the corresponding polynomial

$$\begin{aligned}
p(x) &= x^2 - 4x + 3 \\
&= (x-3) \cdot (x-1)
\end{aligned}$$

the vectors $\vec{z}_i$ are

$$\begin{aligned}
\vec{z}_0 &= \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \\
\vec{z}_1 &= \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix} = (M - (1)I) \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \\
\vec{z}_2 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (M - (3)I) \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix}
\end{aligned}$$

The equation

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (M - (3)I) \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix}$$

rearranged as

$$M \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix} = (3)I \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix}$$

equivalently

$$\begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix} \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix} = 3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix}$$

simplified

$$\begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix} \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix} = 3 \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix}$$

identifies $\begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix}$ as eigenvector with eigenvalue 3 for matrix $\begin{pmatrix} -67 & 116 & 48 \\ -25 & 44 & 18 \\ -35 & 59 & 25 \end{pmatrix}$.

5.4 Computing Eigenvalues and Eigenvectors

The above procedure illustrates that every square matrix has an eigenvector and corresponding eigenvalue. However, the procedure depends on the initial vector, on the order of which polynomial roots are listed in the expansion of the polynomial; it is also tedious. It does indicate that there is a relation between eigenvalues and eigenvectors, evaluating polynomials at matrices and linear dependence.

Once the fact that every square matrix is has at least one eigenvector is established, the next natural step is to find all eigenvectors and eigenvalues.

We can write $A\vec{u} = \lambda\vec{u}$ as $A\vec{u} = \lambda I\vec{u}$ or equivalently

$$(A - \lambda I)\vec{u} = \vec{0}$$

Since we want a non-zero vector $\vec{u}$ the goal reduces to finding non-zero linear combination of the columns of $A - \lambda I$ that evaluate to the zero vector. In other words λ should make the columns of $A - \lambda I$ linearly dependent. Whether $A - \lambda I$ has linearly dependent columns can be checked by looking at its determinant. In other words computing the $\det(A - \lambda I)$ and finding for which λ's the resulting determinant is zero. The determinant of $A - \lambda I$ is a polynomial in λ, so we need the roots of that polynomial. Then for each root we find the non-trivial solutions of $(A - \lambda_i I)\vec{u} = 0$ and obtain the eigenvectors. This argument prove the following,

Theorem 60. *The number λ is an eigenvalue of A if and only if* $\det(A - \lambda I) = 0$*.*

Then an algorithm to find the eigenvectors and eigenvalues for a matrix A proceeds as follows:

1. compute the determinant of $A - \lambda I$
2. compute the roots $\lambda_1, \ldots, \lambda_n$ of the resulting polynomial, the n-roots are the eigenvalues
3. for each eigenvalue i find the corresponding eigenvector by computing $(A - \lambda_i I)\vec{x} = 0$

Definition 56. *For a matrix A the polynomial* $\det(A - \lambda I)$ *is called* the characteristic polynomial *and the equation* $\det(A - \lambda I) = 0$ *is called* the characteristic equation.

5.4.1 Example $2 \times 2 \to 2_1^1, -1_1^1$

Consider matrix A and the corresponding $A - zI$:

$$A = \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} \quad \to \quad A - zI = \begin{pmatrix} -z+4 & -5 \\ 2 & -z-3 \end{pmatrix}$$

The characteristic polynomial is

$$\begin{aligned} p(z) &= det(A - zI) = \det \begin{pmatrix} -z+4 & -5 \\ 2 & -z-3 \end{pmatrix} \\ &= z^2 - z - 2 \\ &= (2 - z) \cdot (-1 - z) \end{aligned}$$

For the solution 2 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned} (A - (2)\, I)\, \vec{x} &= \vec{0} \\ \left(\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} - (2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 2 & -5 \\ 2 & -5 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

The set of solutions is

$$\mathbf{V}_2 = \left\{ s_0 \begin{pmatrix} 5 \\ 2 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

For the solution -1 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (-1)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} - (-1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 5 & -5 \\ 2 & -2 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_{-1} = \left\{ s_0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

5.4.2 Example $3 \times 3 \to 2_1^1, 1_1^1, -3_1^1$

Consider matrix A and the corresponding $A - zI$:

$$A = \begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix} \quad \to \quad A - zI = \begin{pmatrix} -z-9 & 14 & 4 \\ 3 & -z & -2 \\ -18 & 22 & -z+9 \end{pmatrix}$$

The characteristic polynomial is

$$\begin{aligned}
p(z) &= det(A - zI) = \det \begin{pmatrix} -z-9 & 14 & 4 \\ 3 & -z & -2 \\ -18 & 22 & -z+9 \end{pmatrix} \\
&= -z^3 + 7z - 6 \\
&= (2 - z) \cdot (1 - z) \cdot (-3 - z)
\end{aligned}$$

For the solution 2 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (2)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix} - (2) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -11 & 14 & 4 \\ 3 & -2 & -2 \\ -18 & 22 & 7 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_{2} = \left\{ s_0 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

For the solution 1 of the characteristic equation $p(x) = 0$ we have

$$
\begin{aligned}
(A - (1)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix} - (1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -10 & 14 & 4 \\ 3 & -1 & -2 \\ -18 & 22 & 8 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

The set of solutions is

$$
\mathbf{V}_1 = \left\{ s_0 \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}
$$

For the solution -3 of the characteristic equation $p(x) = 0$ we have

$$
\begin{aligned}
(A - (-3)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix} - (-3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -6 & 14 & 4 \\ 3 & 3 & -2 \\ -18 & 22 & 12 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

The set of solutions is

$$
\mathbf{V}_{-3} = \left\{ s_0 \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}
$$

5.4.3 Example $3 \times 3 \to 2_1^2, 0_1^1$

Consider matrix A and the corresponding $A - zI$:

$$
A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{pmatrix} \quad \to \quad A - zI = \begin{pmatrix} -z+1 & 2 & 1 \\ 2 & -z & -2 \\ -1 & 2 & -z+3 \end{pmatrix}
$$

The characteristic polynomial is

$$
\begin{aligned}
p(z) &= det(A - zI) = \det \begin{pmatrix} -z+1 & 2 & 1 \\ 2 & -z & -2 \\ -1 & 2 & -z+3 \end{pmatrix} \\
&= -z^3 + 4z^2 - 4z \\
&= (0 - z) \cdot (2 - z)^2
\end{aligned}
$$

For the solution 0 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (0)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{pmatrix} - (0) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_0 = \left\{ s_0 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

For the solution 2 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (2)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{pmatrix} - (2) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -1 & 2 & 1 \\ 2 & -2 & -2 \\ -1 & 2 & 1 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_2 = \left\{ s_0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

5.4.4 Example $3 \times 3 \to 4_1^3$

Consider matrix A and the corresponding $A - zI$:

$$A = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 1 & -1 \\ -7 & 9 & 9 \end{pmatrix} \quad \to \quad A - zI = \begin{pmatrix} -z+2 & 2 & 2 \\ 2 & -z+1 & -1 \\ -7 & 9 & -z+9 \end{pmatrix}$$

The characteristic polynomial is

$$\begin{aligned}
p(z) &= det(A - zI) = \det \begin{pmatrix} -z+2 & 2 & 2 \\ 2 & -z+1 & -1 \\ -7 & 9 & -z+9 \end{pmatrix} \\
&= -z^3 + 12z^2 - 48z + 64 \\
&= (4 - z)^3
\end{aligned}$$

For the solution 4 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (4)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} 2 & 2 & 2 \\ 2 & 1 & -1 \\ -7 & 9 & 9 \end{pmatrix} - (4) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -2 & 2 & 2 \\ 2 & -3 & -1 \\ -7 & 9 & 5 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_4 = \left\{ s_0 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

5.4.5 Example $3 \times 3 \to 4_2^3$

Consider matrix A and the corresponding $A - zI$:

$$A = \begin{pmatrix} 7 & -1 & -2 \\ 3 & 3 & -2 \\ 3 & -1 & 2 \end{pmatrix} \quad \to \quad A - zI = \begin{pmatrix} -z+7 & -1 & -2 \\ 3 & -z+3 & -2 \\ 3 & -1 & -z+2 \end{pmatrix}$$

The characteristic polynomial is

$$\begin{aligned}
p(z) &= det(A - zI) = \det \begin{pmatrix} -z+7 & -1 & -2 \\ 3 & -z+3 & -2 \\ 3 & -1 & -z+2 \end{pmatrix} \\
&= -z^3 + 12z^2 - 48z + 64 \\
&= (4 - z)^3
\end{aligned}$$

For the solution 4 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (4)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} 7 & -1 & -2 \\ 3 & 3 & -2 \\ 3 & -1 & 2 \end{pmatrix} - (4) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 3 & -1 & -2 \\ 3 & -1 & -2 \\ 3 & -1 & -2 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_4 = \left\{ s_0 \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} + s_1 \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

5.4.6 Example $3 \times 3 \to 4_3^3$

Consider matrix A and the corresponding $A - zI$:

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \to \quad A - zI = \begin{pmatrix} -z+4 & 0 & 0 \\ 0 & -z+4 & 0 \\ 0 & 0 & -z+4 \end{pmatrix}$$

The characteristic polynomial is

$$\begin{aligned} p(z) &= det(A - zI) = \det \begin{pmatrix} -z+4 & 0 & 0 \\ 0 & -z+4 & 0 \\ 0 & 0 & -z+4 \end{pmatrix} \\ &= -z^3 + 12z^2 - 48z + 64 \\ &= (4-z)^3 \end{aligned}$$

For the solution 4 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned} (A - (4)\, I)\, \vec{x} &= \vec{0} \\ \left(\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} - (4) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

The set of solutions is

$$\mathbf{V}_4 = \left\{ s_0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + s_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

5.4.7 Example $5 \times 5 \to 0_2^2, -1_2^3$

Consider matrix A

$$A = \begin{pmatrix} -1 & 0 & 3 & 1 & 5 \\ 0 & -1 & 6 & 2 & 10 \\ 3 & -1 & -4 & -1 & -6 \\ -4 & 3 & -3 & -2 & -7 \\ -1 & 0 & 3 & 1 & 5 \end{pmatrix}$$

and the corresponding $A - zI$

$$A - zI = \begin{pmatrix} -z-1 & 0 & 3 & 1 & 5 \\ 0 & -z-1 & 6 & 2 & 10 \\ 3 & -1 & -z-4 & -1 & -6 \\ -4 & 3 & -3 & -z-2 & -7 \\ -1 & 0 & 3 & 1 & -z+5 \end{pmatrix}$$

The characteristic polynomial is

$$\begin{aligned}
p(z) &= det(A - zI) = \det \begin{pmatrix} -z-1 & 0 & 3 & 1 & 5 \\ 0 & -z-1 & 6 & 2 & 10 \\ 3 & -1 & -z-4 & -1 & -6 \\ -4 & 3 & -3 & -z-2 & -7 \\ -1 & 0 & 3 & 1 & -z+5 \end{pmatrix} \\
&= -z^5 - 3z^4 - 3z^3 - z^2 \\
&= (0-z)^2 \cdot (-1-z)^3
\end{aligned}$$

For the solution 0 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (0)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} -1 & 0 & 3 & 1 & 5 \\ 0 & -1 & 6 & 2 & 10 \\ 3 & -1 & -4 & -1 & -6 \\ -4 & 3 & -3 & -2 & -7 \\ -1 & 0 & 3 & 1 & 5 \end{pmatrix} - (0) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -1 & 0 & 3 & 1 & 5 \\ 0 & -1 & 6 & 2 & 10 \\ 3 & -1 & -4 & -1 & -6 \\ -4 & 3 & -3 & -2 & -7 \\ -1 & 0 & 3 & 1 & 5 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_0 = \left\{ s_0 \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s_1 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ -1 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

For the solution -1 of the characteristic equation $p(x) = 0$ we have

$$\begin{aligned}
(A - (-1)\, I)\, \vec{x} &= \vec{0} \\
\left(\begin{pmatrix} -1 & 0 & 3 & 1 & 5 \\ 0 & -1 & 6 & 2 & 10 \\ 3 & -1 & -4 & -1 & -6 \\ -4 & 3 & -3 & -2 & -7 \\ -1 & 0 & 3 & 1 & 5 \end{pmatrix} - (-1) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right) \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 0 & 0 & 3 & 1 & 5 \\ 0 & 0 & 6 & 2 & 10 \\ 3 & -1 & -3 & -1 & -6 \\ -4 & 3 & -3 & -1 & -7 \\ -1 & 0 & 3 & 1 & 6 \end{pmatrix} \vec{x} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

The set of solutions is

$$\mathbf{V}_{-1} = \left\{ s_0 \begin{pmatrix} 1 \\ 2 \\ 0 \\ -5 \\ 1 \end{pmatrix} + s_1 \begin{pmatrix} 0 \\ 0 \\ 1 \\ -3 \\ 0 \end{pmatrix} \mid \forall i, s_i \in \mathbb{R} \right\}$$

5.4.8 Remarks

In Section 5.3, we established that every matrix has at least one eigenvector. By definition of eigenvalue λ we have that the columns of the matrix $A - \lambda I$ are linearly dependent, thus in this section we establish that for every eigenvalue there is at least one corresponding eigenvector.

The number of eigenvalues is equal to the number of roots of the characteristic polynomial $\det(A - \lambda I)$ which over the complex numbers is the order of the matrix A.

Next we will work on the number of linearly independent eigenvectors. It is worth noting that different eigenvalues have linearly independent eigenvectors. The largest number of linearly independent eigenvector can be less than the order of A. A matrix of order n with n linearly independent eigenvectors can be *diagonalized*.

5.5 Properties of eigenvalues and eigenvectors

Definition 57. *The* eigenspace of a transformation ϕ *associated with the eigenvalue* λ *is*

$$\mathbf{V}_\lambda = \{\vec{\zeta} \mid \phi(\vec{\zeta}) = \lambda\vec{\zeta}\}$$

The eigenspace of a matrix is analogous.

Theorem 61. *An eigenspace is a subspace.*

Proof. The zero vector is in any eigenspace since $A\vec{0} = \lambda\vec{0}$ for any eigenvalue λ. So a subspace is non-empty and it remains to see that an eigenspace is closed under linear combinations. Let $\vec{v}, \vec{u} \in \mathbf{V}_\lambda$; for any constants α and β

$$A(\alpha\vec{v} + \beta\vec{u}) = \alpha A\vec{v} + \beta A\vec{u} = \alpha\lambda\vec{v} + \beta\lambda\vec{u} = \lambda(\alpha\vec{v} + \beta\vec{u})$$

and therefore $\alpha\vec{v} + \beta\vec{u} \in \mathbf{V}_\lambda$. By Theorem 23 the result follows. $\square$

With the above result in mind

Definition 58 (multiplicity)**.** *Let A be a square matrix of order n with characteristic polynomial:*

$$\det(A - \lambda I) \quad = \quad (\lambda_0 - \lambda)^{m_0}(\lambda_1 - \lambda)^{m_1} \cdots (\lambda_t - \lambda)^{m_t}$$

The algebraic multiplicity *of $\lambda_i \in \{\lambda_0 \dots \lambda_t\}$ is m_i. The* geometric multiplicity *of the eigenvalue λ_i is the dimension of the corresponding eigenspace $\mathbf{V}_{\lambda_i}$.*

Example: From the examples above

	eigenvalue	algebraic multiplicity	geometric multiplicity
§5.4.1	2	1	1
	-1	1	1
§5.4.2	2	1	1
	1	1	1
	-3	1	1
§5.4.3	2	2	1
	0	1	1
§5.4.4	4	3	1
§5.4.5	4	3	2
§5.4.6	4	3	3
§5.4.7	0	2	2
	-1	3	2

The geometric multiplicity of an eigenvalue is no larger than its algebraic multiplicity (while true this claim is not proved here), and is at least one (since every eigenvalue has at least one non-trivial eigenvector).

Theorem 62. *A set of eigenvectors of corresponding to distinct eigenvalues is linearly independent.*

Proof. We will use induction on the number of eigenvalues. The base step is that there are zero eigenvalues. Then the set of associated vectors is empty and so is linearly independent.

For the inductive step assume that the statement is true for any set of $k \geq 0$ distinct eigenvalues. Consider distinct eigenvalues $\lambda_1, \ldots, \lambda_{k+1}$ and let $\vec{v}_1, \ldots, \vec{v}_{k+1}$ be associated eigenvectors. Suppose that

$$\vec{0} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1}.$$

Derive two equations from that, the first by multiplying by λ_{k+1} on both sides

$$\vec{0} = c_1\lambda_{k+1}\vec{v}_1 + \cdots + c_{k+1}\lambda_{k+1}\vec{v}_{k+1}$$

and the second by applying the map to both sides

$$\vec{0} = c_1t(\vec{v}_1) + \cdots + c_{k+1}t(\vec{v}_{k+1}) = c_1\lambda_1\vec{v}_1 + \cdots + c_{k+1}\lambda_{k+1}\vec{v}_{k+1}$$

(applying the matrix gives the same result). Subtract the second from the first.

$$\vec{0} = c_1(\lambda_{k+1} - \lambda_1)\vec{v}_1 + \cdots + c_k(\lambda_{k+1} - \lambda_k)\vec{v}_k + c_{k+1}(\lambda_{k+1} - \lambda_{k+1})\vec{v}_{k+1}$$

The $\vec{v}_{k+1}$ term vanishes. Then the induction hypothesis gives that

$$c_1(\lambda_{k+1} - \lambda_1) = 0 \ldots c_k(\lambda_{k+1} - \lambda_k) = 0$$

The eigenvalues are distinct so the coefficients $c_1, \ldots, c_k$ are all 0. With that we are left with the equation $\vec{0} = c_{k+1}\vec{v}_{k+1}$ so c_{k+1} is also 0. $\square$

5.6 Diagonal form of a matrix

5.6.1 Similarity

Definition 59. *Two matrices A and B are called* similar *if there is an invertible matrix S such that $A = S^{-1}BS$.*

From $\left(S^{-1}\right)^{-1} = S$ if A is similar to B then B is similar to A.

Example: from

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

we conclude $\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$ is similar to $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Definition 60. *A matrix A is* diagonalizable *if it is similar to a diagonal matrix D.*

5.6.2 Example

The matrix from §5.4.2 has three linearly independent eigenvectors vectors

$$\vec{s}_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \quad \vec{s}_2 = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} \quad \vec{s}_3 = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$$

Construct a matrix

$$S = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \vec{s}_1 & \vec{s}_2 & \vec{s}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}$$

Since the columns of S are linearly independent the matrix is invertible with inverse

$$S^{-1} = \begin{pmatrix} 3 & -1 & -2 \\ 2 & -2 & -1 \\ -3 & 2 & 2 \end{pmatrix}$$

Consider $R = AS$

$$\underbrace{\begin{pmatrix} 4 & -6 & 3 \\ 2 & 0 & 1 \\ 4 & -9 & 4 \end{pmatrix}}_{R} = \overbrace{\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}}^{S}$$

Let

$$\vec{r}_1 = \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix} \quad \vec{r}_2 = \begin{pmatrix} -6 \\ 0 \\ -9 \end{pmatrix} \quad \vec{r}_3 = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$$

We have

$$\overbrace{\begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix}}^{\vec{r}_1} = \overbrace{\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}}^{\vec{s}_1} = \overbrace{2}^{d_{11}} \overbrace{\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}}^{\vec{s}_1}$$

$$\overbrace{\begin{pmatrix} -6 \\ 0 \\ -9 \end{pmatrix}}^{\vec{r}_2} = \overbrace{\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}}^{\vec{s}_2} = \overbrace{-3}^{d_{22}} \overbrace{\begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}}^{\vec{s}_2}$$

$$\overbrace{\begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}}^{\vec{r}_3} = \overbrace{\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}}^{\vec{s}_3} = \overbrace{1}^{d_{33}} \overbrace{\begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}}^{\vec{s}_3}$$

where d_{ii} is the eigenvalue corresponding to eigenvector $\vec{s}_i$. For every vector $\vec{s}_i$ we have

$$\begin{aligned} \vec{s}_1 &= \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ \vec{s}_2 &= \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ \vec{s}_3 &= \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

Adding the eigenvalues and using the fact that multiplication with a constant commutes with vector multiplication we get

$$\begin{aligned}
d_{11}\vec{s}_1 &= 2\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\left[2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \\
d_{22}\vec{s}_2 &= -3\begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = -3\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\left[-3\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ -3 \\ 0 \end{pmatrix} \\
d_{33}\vec{s}_3 &= 1\begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = 1\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\left[1\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

So for every i we have $\vec{r}_i = S(d_{ii}\vec{e}_i) = S\vec{d}_i$, where

$$\begin{aligned}
\vec{d}_1 &= d_{11}\vec{e}_1 = 2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \\
\vec{d}_2 &= d_{22}\vec{e}_2 = -3\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -3 \\ 0 \end{pmatrix} \\
\vec{d}_3 &= d_{33}\vec{e}_3 = 1\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

Combine the vectors $\vec{d}_i$ into a (diagonal) matrix

$$D = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \vec{d}_1 & \vec{d}_2 & \vec{d}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

to obtain

$$\begin{pmatrix} 4 & -6 & 3 \\ 2 & 0 & 1 \\ 4 & -9 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus $R = AS$ and $R = SD$

$$\underbrace{\begin{pmatrix} 4 & -6 & 3 \\ 2 & 0 & 1 \\ 4 & -9 & 4 \end{pmatrix}}_{R} = \overbrace{\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}}^{S}$$

$$\overbrace{\begin{pmatrix} 4 & -6 & 3 \\ 2 & 0 & 1 \\ 4 & -9 & 4 \end{pmatrix}}^{R} = \underbrace{\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}}_{S} \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{D}$$

from which one concludes $AS = SD$. Multiplyting on the right with S^{-1} we get

$$\overbrace{\begin{pmatrix} -9 & 14 & 4 \\ 3 & 0 & -2 \\ -18 & 22 & 9 \end{pmatrix}}^{A} = \overbrace{\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}}^{S} \overbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{D} \overbrace{\begin{pmatrix} 3 & -1 & -2 \\ 2 & -2 & -1 \\ -3 & 2 & 2 \end{pmatrix}}^{S^{-1}}$$

which shows the matrix A is diagonalizable.

5.6.3 Example

From $A = SDS^{-1}$ where

$$\overbrace{\begin{pmatrix} 8 & -6 & -12 \\ 3 & -1 & -6 \\ 3 & -3 & -4 \end{pmatrix}}^{A} = \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \overbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}}^{D} \overbrace{\begin{pmatrix} 1 & 0 & -2 \\ 0 & -1 & 1 \\ -1 & 1 & 2 \end{pmatrix}}^{S^{-1}}$$

the matrix A is diagonalizable. Multiplying on both sides from the right with S we get $AS = SD$; suppose the result of the multiplication is R so

$$R = AS \qquad \text{and} \qquad R = SD$$

that is

$$\underbrace{\begin{pmatrix} 6 & 4 & -2 \\ 2 & 0 & -1 \\ 2 & 2 & -1 \end{pmatrix}}_{R} = \overbrace{\begin{pmatrix} 8 & -6 & -12 \\ 3 & -1 & -6 \\ 3 & -3 & -4 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S}$$

$$\overbrace{\begin{pmatrix} 6 & 4 & -2 \\ 2 & 0 & -1 \\ 2 & 2 & -1 \end{pmatrix}}^{R} = \underbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_{S} \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{D}$$

Let $\vec{r}_1 = \begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix} \quad \vec{r}_2 = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix} \quad \vec{r}_3 = \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}$ that is

$$\vec{r}_1 = \begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix} \quad \vec{r}_2 = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix} \quad \vec{r}_3 = \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}$$

and $S = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \vec{s}_1 & \vec{s}_2 & \vec{s}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ that is

$$\vec{s}_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \vec{s}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \quad \vec{s}_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

From $R = AS$ using properties of matrix multiplication we get $\vec{r}_i = A\vec{s}_i$

$$\overbrace{\begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix}}^{\vec{r}_1} = \overbrace{\begin{pmatrix} 8 & -6 & -12 \\ 3 & -1 & -6 \\ 3 & -3 & -4 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}}^{\vec{s}_1}$$

$$\overbrace{\begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}}^{\vec{r}_2} = \overbrace{\begin{pmatrix} 8 & -6 & -12 \\ 3 & -1 & -6 \\ 3 & -3 & -4 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}}^{\vec{s}_2}$$

$$\overbrace{\begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}}^{\vec{r}_3} = \overbrace{\begin{pmatrix} 8 & -6 & -12 \\ 3 & -1 & -6 \\ 3 & -3 & -4 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}}^{\vec{s}_3}$$

Likewise for $D = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \vec{d}_1 & \vec{d}_2 & \vec{d}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ where

$$\vec{d}_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \quad \vec{d}_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \quad \vec{d}_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

we have

$$
\begin{aligned}
\vec{d}_1 &= \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = d_{11}\vec{e}_1 \\
\vec{d}_2 &= \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = d_{22}\vec{e}_2 \\
\vec{d}_3 &= \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = -1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = d_{33}\vec{e}_3 \\
\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} &= \left(\underbrace{2}_{d_{11}} \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\vec{e}_1}, \quad \underbrace{2}_{d_{22}} \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\vec{e}_2}, \quad \underbrace{-1}_{d_{33}} \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\vec{e}_3} \right)
\end{aligned}
$$

From $R = SD$ using properties of matrix multiplication we obtain $\vec{r}_i = d_{ii}\vec{s}_i$

$$\overbrace{\begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix}}^{\vec{r}_1} = \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \overbrace{\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}}^{\vec{d}_1} = \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \underbrace{(2)}_{d_{11}} \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\vec{e}_1}$$

$$= \underbrace{(2)}_{d_{11}} \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\vec{e}_1} = \underbrace{(2)}_{d_{11}} \overbrace{\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}}^{\vec{s}_1}$$

$$\overbrace{\begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}}^{\vec{r}_2} = \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \overbrace{\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}}^{\vec{d}_2} = \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \underbrace{(2)}_{d_{22}} \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\vec{e}_2}$$

$$= \underbrace{(2)}_{d_{22}} \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\vec{e}_2} = \underbrace{(2)}_{d_{22}} \overbrace{\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}}^{\vec{s}_2}$$

$$\overbrace{\begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}}^{\vec{r}_3} = \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \overbrace{\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}}^{\vec{d}_3} = \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \underbrace{(-1)}_{d_{33}} \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\vec{e}_3}$$

$$= \underbrace{(-1)}_{d_{33}} \overbrace{\begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{S} \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\vec{e}_3} = \underbrace{(-1)}_{d_{33}} \overbrace{\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}}^{\vec{s}_3}$$

As a result we have

$$\vec{r}_i = A\vec{s}_i = d_{ii}\vec{s}_i$$

and by definition we get $\vec{s}_i$ is eigenvector with eigenvalue d_{ii} for matrix A. Since S is invertible its columns are linearly independent therefore $\vec{s}_1, \vec{s}_2, \vec{s}_3$ are linearly independent and therefore we have three linearly independent eigenvectors for matrix A.

5.6.4 Diagonalization

In the above two examples do not depend on the actual values of the constants. The idea can be generalized to the following

Theorem 63. *An $n \times n$ matrix is diagonalizable if and only if it has n linearly independent eigenvectors.*

Proof. Assume $S^{-1}AS = D = (d_1\vec{e}_1, d_2\vec{e}_2, \ldots, d_n\vec{e}_n)$ and let

$$S = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{s}_1 & \vec{s}_2 & & \vec{s}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

alternatively $S\vec{e}_i = \vec{s}_i$. We have that

$$\begin{aligned} (A\vec{s}_1, A\vec{s}_2, \ldots, A\vec{s}_n) &= A(\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_n) \\ &= IAS = SS^{-1}AS = SD = S(d_1\vec{e}_1, d_2\vec{e}_2, \ldots, d_n\vec{e}_n) \\ &= (Sd_1\vec{e}_1, Sd_2\vec{e}_2, \ldots, Sd_n\vec{e}_n) \\ &= (d_1S\vec{e}_1, d_2S\vec{e}_2, \ldots, d_nS\vec{e}_n) \\ &= (d_1\vec{s}_1, d_2\vec{s}_2, \ldots, d_n\vec{s}_n) \end{aligned}$$

Column-wise we have $A\vec{s}_i = d_i\vec{s}_i$; the columns of S are all non-zero since S is invertible and therefore every column of S is an eigenvector for A.

Conversely, assume that A has n linearly independent eigenvectors say $\vec{s}_1, \ldots, \vec{s}_n$ with corresponding eigenvalues $\lambda_1, \lambda_2 \ldots, \lambda_n$ construct a matrix S whose columns are $\vec{s}_1, \ldots, \vec{s}_n$ so

$$S = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{s}_1 & \vec{s}_2 & \ldots & \vec{s}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

alternatively

$$S\vec{e}_i = \vec{s}_i.$$

We have that

$$\begin{aligned} S^{-1}AS &= S^{-1}A(\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_n) \\ &= S^{-1}(A\vec{s}_1, A\vec{s}_2, \ldots, A\vec{s}_n) \\ &= S^{-1}(\lambda_1\vec{s}_1, \lambda_2\vec{s}_2, \ldots, \lambda_n\vec{s}_n) \\ &= S^{-1}(\lambda_1S\vec{e}_1, \lambda_2S\vec{e}_2, \ldots, \lambda_nS\vec{e}_n) \\ &= S^{-1}S(\lambda_1\vec{e}_1, \lambda_2\vec{e}_2, \ldots, \lambda_n\vec{e}_n) \\ &= (\lambda_1\vec{e}_1, \lambda_2\vec{e}_2, \ldots, \lambda_n\vec{e}_n) \\ &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = D \end{aligned}$$

where D is a diagonal matrix with diagonal entries the eigenvalues of A. □

Example: we can have different matrices S by picking different eigenvectors as long as they are linearly independent. Likewise if we reorder the eigenvectors we will get a different (but similar) diagonal matrix.

$$\begin{aligned}
\begin{pmatrix} 8 & -6 & -12 \\ 3 & -1 & -6 \\ 3 & -3 & -4 \end{pmatrix} &= \begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & -2 \\ 0 & -1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \\
&= \begin{pmatrix} -\frac{2}{3} & \frac{9}{2} & 4 \\ 0 & \frac{3}{2} & 2 \\ -\frac{1}{3} & \frac{3}{2} & 2 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 3 & -3 \\ \frac{2}{3} & 0 & -\frac{4}{3} \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}
\end{aligned}$$

Determinant. A matrix A is *invertible* if none of its eigenvalues is zero:

$$\det(A) = \det(S^{-1}\Lambda S) = \det(S^{-1})\det(\Lambda)\det(S) = \det(\Lambda)$$

Diagonalization. A matrix A is *diagonalizable* if it has sufficiently many linearly independent eigenvectors. For example let's try to diagonalize $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$; the matrix A has only eigenvalue 1 and a single linearly independent eigenvector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. If A is diagonalizable then there is an invertible matrix $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with inverse $S^{-1} = \frac{1}{ad-cb}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Such that

$$\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} = \frac{1}{ad-cb}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

equivalently

$$\begin{aligned}
(ad-cb)\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} &= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
&= \begin{pmatrix} d-b & -b \\ -c+a & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
&= \begin{pmatrix} a(d-b)-bc & bd-b^2-bd \\ -ac+a^2+ac & (a-c)b+ad \end{pmatrix} \\
&= \begin{pmatrix} a(d-b)-bc & -b^2 \\ a^2 & (a-c)+ad \end{pmatrix}
\end{aligned}$$

From here

$$\begin{aligned}
(ad-cb)x &= a(d-b)-bc \\
0 &= -b^2 \\
0 &= a^2 \\
(ad-cb)y &= (a-c)b+ad
\end{aligned}$$

concluding that $a = b = 0$. In that case for the matrix S we have $S = \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}$ which is not an invertible matrix, a contradiction with S being invertible.

Definition 61 (defective matrix)**.** *A matrix A that is not diagonalizable is call* defective*.*