Chapter 1

Linear Systems

1.1 Linearity

Definition 1.1.1 (linear combination)**.** *A* linear combination *of $x_1 \dots x_n$ 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 $\mathbb{R}$ or $\mathbb{C}$.*

Examples:

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

Definition 1.1.2 (linear equation)**.** *A* linear equation *in the set of variables X, where without loss of generality $X = \{x_1 \dots x_n\}$ is an equation of the form*

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

The value d is the constant *of the linear equation and similar to the coefficients it belongs to the real (complex) numbers.*

Often linear equation are denoted via the summation symbol $\sum$ as follows

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

Examples:

- These are linear equations

  $2x_1 + 3x_2 + x_3 = 5$

$$x_1 + x_2 + x_3 + x_4 = 8$$

- These are *not* linear equations

$$\sqrt{x_1} + 3x_2 + x_4 = 3$$
$$x_1x_2 + x_2 + x_4 = 6$$

1.2 Solutions to a system of linear equations

Definition 1.2.1 (solution of an equation)**.** *An n-tuple $s_1, \ldots, s_n \in \mathbb{R}$*[1] *is a solution to the linear equation*

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

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

Example: $1, 0, -3)$ is solutions to $x_1 + 2x_2 + x_3 = 1$, but $(-3, 0, 1)$ is not a solution. 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 $x_1 + x_2 + x_3 = 3$. In general for the latter we have $x_3 = t, x_2 = s$ and $x_1 = 3 - t - s$ where $s, t \in \mathbb{C}$.

Definition 1.2.2 (system of linear equations)**.** *A system of linear equation is a set of linear equations in the same set of variables $x_1, \ldots, x_n$:*

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

Definition 1.2.3 (solution to a system of linear equations)**.** *An n-tuple $(s_1, \ldots, s_n) \in \mathbb{R}^n$ is a solution to a linear system of equations if it is solution for each equation.*

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

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

[1] Order is important

Example: For arbitrary real (complex) values s and t

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

is a solution to

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

1.3 Equivalent systems

Definition 1.3.1. *Two systems of equations S_1 and S_2 are said to be equivalent if they have the same set of solutions.*

Example: For example

$$\begin{aligned} x_1 - 2x_2 + 3x_3 + x_4 &= -3 \\ 2x_1 - x_2 + 3x_3 - x_4 &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x_1 - 2x_2 + 3x_3 + x_4 &= -3 \\ 4x_1 - 5x_2 + 9x_3 + x_4 &= -6 \end{aligned}$$

both have the same set of solutions:

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

Here $s, t \in \mathbb{R}$. However, the following system while closely related is not equivalent as it has a unique solution.

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

Consider the following two systems of linear equations:

$$\begin{aligned} x_1 - 2x_2 + 3x_3 + x_4 &= -3 \\ 4x_1 - 5x_2 + 9x_3 + x_4 &= -6 \end{aligned} \qquad \begin{aligned} y_1 - 2y_2 + 3y_3 + y_4 &= -3 \\ 4y_1 - 5y_2 + 9y_3 + y_4 &= -6 \end{aligned}$$

The former one is in indeterminates $\{x_1, x_2, x_3, x_4\}$ and the latter in indeterminates $\{y_1, y_2, y_3, y_4\}$. Those are trivially equivalent and therefore we may conclude that the actual information 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. This gives a way to represent system of linear equations via so called *matrices*.

1.4 Matrices and vectors

Definition 1.4.1 (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 following is a 3×5 matrix: $\begin{pmatrix} 1 & 3 & 3 & 2 & 1 \\ 3 & 9 & -6 & 4 & 3 \\ 2 & 6 & -4 & 2 & 2 \end{pmatrix}$, but 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 1.4.2 (equal matrices). *Let*

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

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 $a_{ij} = b_{ij}$

Definition 1.4.3 (square matrix). *An $m \times n$ matrix is called* square of order m *if $m = n$.*

Definition 1.4.4 (row vector). *An $1 \times n$ matrix is called a* row vector*. The entries in a vector are also called* components*.*

Definition 1.4.5 (column vector). *An $m \times 1$ matrix is called a* column vector.[2]

1.5 Representations of system of linear equations

Definition 1.5.1 (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*

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

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 $j \le n$ and b_i otherwise. That is*

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

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

Example: Consider 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 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

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

Example: another example

$$\begin{array}{rrrrrcr} 2y_1 & +6y_2 & -4y_3 & +2y_4 & +2y_5 & = & -2 \\ & & & y_4 & & = & -1 \\ & & 5y_3 & +y_4 & & = & 9 \end{array} \quad \rightarrow \quad \left(\begin{array}{rrrrr|r} 2 & 6 & -4 & 2 & 2 & -2 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 5 & 1 & 0 & 9 \end{array}\right)$$

Definition 1.5.2 (vector representation of a system of linear equations)**.** *Let S be a system of linear equations in $\{x_1, \ldots, x_n\}$ given by*

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

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

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

1.6 Echelon form and back substitution

Definition 1.6.1 (leading variable)**.** *In each row of a system, the first variable with a nonzero coefficient is the rows* leading variable*.*

Definition 1.6.2 (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.*

Examples:

$$\begin{array}{rrrrcr} x & +3y & +3z & +2w & = & 1 \\ 2x & +6y & +9z & +5w & = & 3 \\ -1x & -3y & +3z & & = & 1 \end{array} \qquad \begin{array}{rrrrcr} x & +3y & +3z & +2w & = & 1 \\ & & 3z & +w & = & 1 \\ & & & 0 & = & 0 \end{array}$$

The first system is in Echelon form. Overall it has only one leading variable – x. In the right system, which happens to be in Echelon form x and z are the leading variables. There are no other leading variables.

Definition 1.6.3 (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 $s \in \mathbb{R}$. Then

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

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

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

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

$$x_3 \quad = \quad -2 - \frac{2}{7} + \frac{4}{7}s - 2s = \frac{16}{7} - \frac{10}{7}s$$

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

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 t. Note that the values of t and s are independent from each other, so now we have

$$\begin{aligned}
x_2 &= t \\
x_3 &= -2 - \frac{2}{7} + \frac{4}{7}s - 2s = \frac{16}{7} - \frac{10}{7}s \\
x_4 &= \frac{1}{7} - \frac{2}{7}s \\
x_5 &= 2s \\
x_6 &= s;
\end{aligned}$$

and with this information we can also express x_1 as

$$\begin{aligned}
2x_1 &= 3 - 2t - 3\left(\frac{16}{7} - \frac{10}{7}s\right) - 3\left(\frac{1}{7} - \frac{2}{7}s\right) - s \\
&= -\frac{30}{7} - 2t + \frac{29}{7}s
\end{aligned}$$

to get the complete set of solutions as

$$\begin{aligned}
x_1 &= -\frac{30}{14} - t + \frac{29}{14}s \\
x_2 &= t \\
x_3 &= -2 - \frac{2}{7} + \frac{4}{7}s - 2s = \frac{16}{7} - \frac{10}{7}s \\
x_4 &= \frac{1}{7} - \frac{2}{7}s \\
x_5 &= 2s \\
x_6 &= s;
\end{aligned}$$

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

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

has no solution (the solution set is empty) since the last equation has no solution. In this case we say the system is *inconsistent*.

Definition 1.6.4 (inconsistent constent). *A linear system of equations is said to be* inconsistent *if it has no solutions; otherwise they are called* consistent.

Example: Here is another example of a system that can easily be identified as inconsistent. Note that a system being consistent or inconsistent is independent from whether it is in Echelon form or not.

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

Observe that even though the system given by

$$\left(\begin{array}{cccccccccc|c} 2 & 2 & 3 & 3 & 0 & 1 & 7 & 9 & 1 & 11 & 1 \\ 0 & 5 & 3 & 5 & 0 & 1 & 7 & 9 & 12 & 76 & 2 \\ 5 & 2 & 9 & 3 & 5 & 6 & 7 & 14 & -12 & 1 & 3 \\ 2 & 0 & 5 & 3 & 0 & 1 & 7 & 9 & 7 & 15 & 4 \\ 1 & 2 & 3 & 7 & 1 & 1 & 6 & 9 & 12 & 16 & 5 \\ 2 & 0 & 3 & 3 & 0 & 1 & 7 & 9 & 1 & 10 & 6 \end{array}\right)$$

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

$$\left(\begin{array}{cccccccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & \frac{30}{37} & \frac{40}{37} & -\frac{293}{74} & -\frac{4195}{148} & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 3 & \frac{5}{2} & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & \frac{30}{37} & \frac{40}{37} & \frac{57}{37} & \frac{126}{37} & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & -\frac{127}{37} & -\frac{120}{37} & \frac{65}{74} & -\frac{5471}{148} & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{109}{37} & \frac{133}{37} & -\frac{174}{37} & \frac{1812}{37} & 6 \end{array}\right).$$

The reason is that there is no general procedure that can apply back substitution to the second system but not for the first. Consequently, to solve systems of linear equation we need a method to transform them 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

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

Overall we obtain a new system of linear equations given by

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

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

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

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:

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

Observe that the row 1 of the right side is $2 * Eqn1 + 0 * Eqn2 + 1 * Eqn3$ 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 1.7.1 (scalar matrix multiplication)**.** *Let $A = \{a_{ij}\}$ be a $m \times n$ matrix and r be a real number. We define rA as the $m \times n$ with entries $\{a_{ij}\}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.*

When adding two Equations we added constants in front of variables, which in vector forms is simple component-wise addition (not the row vectors must have the same number of components). Again generalizing to matrix addition we get

Definition 1.7.2 (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 $1 \leq j \leq n$.*

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

Definition 1.7.3 (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 $c_{uv} = \sum_{i=1}^{n} a_{ui} b_{iv}$.

Definition 1.7.4 (linear combination of vectors)**.** *A* linear combination *of vectors $\vec{v}_1 \dots \vec{v}_n$ 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 $\mathbb{R}$.

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.

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. In particular we can write from the vector form of a system of linear equation we can write the matrix form of a system of linear equations $A\vec{x} = \vec{b}$ where A is the coefficient matrix of the system.

Matrix multiplication and 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:

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

Furthermore there is a *zero matrix* O, whose entries are all zeroes such that $A + O = O + A = A$. Similarly, for every matrix A there is a matrix B such that $A + B = O$. 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. Transpose of a matrix is

Definition 1.7.5 (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 $1 \leq u \leq n$ and $1 \leq v \leq m$.*

Before listing the properties of multiplication we make the following definitions:

Definition 1.7.6 (lower 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* lower triangular *matrix.*

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

Definition 1.7.7 (upper 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* upper triangular *matrix.*

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

Definition 1.7.8 (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.*

Definition 1.7.9 (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.*

Definition 1.7.10 (identity matrix)**.** *The scalar matrix with diagonal entries equal to* 1 *is called the* identity *matrix and denoted by I.*

Definition 1.7.11 (square matrix)**.** *An $n \times n$ matrix is called a* square *matrix. Often it is called* square matrix of order *n.*

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
- How about zero and one? For square matrices defined the identity matrix
- define diagonal matrix by motivating is as a generalization of I in the sense it is a matrix that commutes with any other matrix.
- Looking at $ax = b$ and the solution $x = \frac{b}{a}$ extend to inverse of a matrix. The solution should be $x = \frac{b}{A}$.
- Observe that for one variable one equation we have $0x = b$ not solvable for $b \neq 0$. This follows from the fact that the solution is $\frac{1}{a}ax = \frac{1}{a}b$ is not possible if $a = 0$. Naturally this generalizes to matrices. Motivate zero divisors with $0x = b$ and the impossibility of division.

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 $AB = 0$ or $BA = 0$ then A is called a *zero divisor*.

Example Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Both A and B are zero divisors.

Example

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

Both A and B are zero divisors.

Definition 1.7.12 (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}.*

Theorem 1.7.1. *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 1.7.2. *If A is invertible matrix then the inverse of A is unique.*

Proof. Suppose C and D are both inverses of A by the above Theorem 1.7.1 we have that the left and the right inverse of a matrix are equal. Since D is an inverse of A D is also a left inverse of A and therefore it equals the right inverse of A which is C. Thus $C = D$ completing the argument. □

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 1.8.1. *[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

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

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

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

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

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

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 $b'_j = b_j + cb_kk$ implying $b_j = b'_j - cb_t$. For Equation 1.2 we then have

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

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

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

□

Definition 1.8.1 (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 1.8.1 all the system of equations are equivalent (that is they have the same set of solutions).

Example: of Gauss' method recall the system

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

in matrix form

$$S : \left(\begin{array}{rrrrr|r}
1 & 3 & 3 & 2 & 1 & 7 \\
3 & 9 & -6 & 4 & 3 & -7 \\
2 & 6 & -4 & 2 & 2 & -4
\end{array}\right)$$

$$S1:\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 3 & 9 & -6 & 4 & 3 & -7 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right)$$

$$S2:\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & -10 & -2 & 0 & -18 \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{array}\right)\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 2 & 6 & -4 & 2 & 2 & -4 \end{array}\right)$$

$$S3:\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & -\frac{2}{3} & 0 & \frac{2}{3} \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{array}\right)\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & -10 & -2 & 0 & -18 \end{array}\right)$$

which has solution set given by

$$x_1 = 3 - 3t - s, x_2 = t, x_3 = 2, x_4 = -1, x_5 = s$$

in vector form

$$\left\{\left(\begin{array}{c} 3 \\ 0 \\ 2 \\ -1 \\ 0 \end{array}\right) + \left(\begin{array}{c} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{array}\right)s + \left(\begin{array}{c} 0 \\ 1 \\ 0 \\ 0 \\ -3 \end{array}\right)t \mid s, t \in \mathbb{C}\right\}$$

We can continue with elementary row operations to obtain the *Reduced Echelon form*

$$S4:\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{3}{2} \end{array}\right)\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & -\frac{2}{3} & 0 & \frac{2}{3} \end{array}\right)$$

$$S5:\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & 0 & 0 & -30 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & -2 & 0 & -28 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right)$$

$$S6:\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & -\frac{1}{15} & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & -15 & 0 & 0 & -30 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right)$$

$$S7:\left(\begin{array}{ccccc|c} 1 & 3 & 0 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right) = \left(\begin{array}{ccc} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{ccccc|c} 1 & 3 & 3 & 2 & 1 & 7 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right)$$

With the reduced Echelon form it is even easier to identify the solution. It is especially useful if you need to solve multiple SLEs with the same coefficient matrix.

Example: SLE with no solutions

$$S1:\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 5 & 3 \\ -1 & -3 & 3 & 0 & 2 \end{array}\right)$$

$$S2:\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 6 & 2 & 3 \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right)\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 5 & 3 \\ -1 & -3 & 3 & 0 & 2 \end{array}\right)$$

$$S3:\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 6 & 2 & 3 \end{array}\right)$$

We can obtain the reduced form

$$S5:\left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

but we already know there is no solution. This can be useful only if we expect to solve a system with the same matrix.

Example: The above equation but different constants

$$S1:\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 5 & 3 \\ -1 & -3 & 3 & 0 & 1 \end{array}\right)$$

$$S2:\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 6 & 2 & 2 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 5 & 3 \\ -1 & -3 & 3 & 0 & 1 \end{array}\right)$$

$$S3:\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 6 & 2 & 2 \end{array}\right)$$

The solution set is $\{(-3t-3s, t, \frac{1}{3}-\frac{1}{3}s, s) \mid s,t \in \mathbb{C}\}$, in vector form:

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ \frac{1}{3} \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix} s + \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} t \mid s,t \in \mathbb{C} \right\}$$

We can obtain the *Reduced Echelon form*

$$S4:\left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$S5:\left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Using the reduced Echelon form it is even easier to identify the solution.

Example: SLE with unique solution (in fact three of these):

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

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

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

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

We can now solve each of those SLE to obtain column wise $\begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ -1 & 3 & 1 \end{pmatrix}$.

1.8.1 Reduced Echelon form and inverse

We can however proceed (see previous example) with Gaussian eliminations to get something even better

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

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

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

$$\left(\begin{array}{rrr|rrr} 5 & 0 & 0 & 10 & 5 & 15 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} 1 & 8 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{rrr|rrr} 5 & -\frac{48}{5} & 0 & \frac{2}{5} & -\frac{23}{5} & -\frac{21}{5} \\ 0 & -\frac{1}{5} & 0 & -\frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 2 & 1 & 3 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right) = \begin{pmatrix} \frac{1}{5} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{rrr|rrr} 1 & -\frac{48}{25} & 0 & \frac{2}{25} & -\frac{23}{25} & -\frac{21}{25} \\ 0 & -\frac{1}{5} & 0 & -\frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -1 & 3 & 1 \end{array}\right)$$

Ultimately we have that

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

Multiplying together the elementary operations, which is possible by associativity of matrix multiplication we have that

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

This is an algorithm to compute the inverse of a matrix if one exists.

1.9 Homogeneous and particular solutions

Example: Recall the system with augmented matrix

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

in Reduced Echelon form we had

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

and solution in vector form

$$\left\{ \begin{pmatrix} 3 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} s + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -3 \end{pmatrix} t \mid s, t \in \mathbb{C} \right\}$$

The system in matrix form is written as

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

Often we will simply write

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

Particular solution. Given the vector addition and scalar multiplication from the previous section take the vector

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

and substitute it in the vector equation of the system. The result is the vector $\begin{pmatrix} 7 \\ -7 \\ -4 \end{pmatrix}$.

Homogeneous part. Repeat the same for the vectors $\begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, the result is $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

Definition 1.9.1 (homogeneous equation)**.** *An equation* $\sum_{i=1}^{n} a_i x_i = b$ *is called* homogeneous *if* $b = 0$.

Definition 1.9.2 (homogeneous system of equations)**.** *A system of equations is called* homogeneous *if each equation is homogeneous.*

In other words in the vector representation of the solution if a vector s satisfies $As = 0$ then it is part of the homogeneous solution and if s satisfies $As = b$ 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.

Theorem 1.9.1. *Any linear system's solution set has the form*

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

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.*

Lemma 1.9.1. *For any homogeneous linear system there exist vectors* $\vec{\beta}_1, \ldots, \vec{\beta}_k$ *such that the solution set of the system is*

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

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. There may be some $0 = 0$ equations; we ignore these (if the system consists only of $0 = 0$ equations then the lemma is trivially true because there are no leading variables). But because the system is homogeneous there are no contradictory equations.

We will use induction to verify that each leading variable can be expressed in terms of free variables. That will finish the proof because we can use the free variables 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 $a_{m,\ell_m} \neq 0$. (The 'ℓ' means "leading" so that x_{ℓ_m} is the leading variable in row m.) This is the bottom row so any variables after the leading one must be free. Move these to the right hand side and divide by a_{m,ℓ_m}

$$x_{\ell_m} = (-a_{m,\ell_m+1}/a_{m,\ell_m})x_{\ell_m+1} + \cdots + (-a_{m,n}/a_{m,\ell_m})x_n \tag{1.6}$$

to express the leading variable in terms of free variables. (There is a tricky technical point here: if in the bottom Equation 1.5 there are no variables to the right of x_{l_m} then $x_{\ell_m} = 0$. This satisfies the statement we are verifying because, as alluded to at the start of this subsection, it has x_{ℓ_m} written as a sum of a number of the free variables, namely as the sum of zero many, under the convention that a trivial sum totals to 0.)

For the inductive step assume that the statement holds for the bottom-most t rows, with $0 \leq t < m-1$. That is, assume that for the m-th equation, and the $(m-1)$-th equation, etc., up to and including the $(m-t)$-th equation, we can express the leading variable in terms of free ones. We must verify that this then also holds for the next equation up, the $(m-(t+1))$-th equation. For that, take each variable that leads in a lower equation $x_{\ell_m}, \ldots, x_{\ell_{m-t}}$ and substitute its expression in terms of free variables. We only need expressions for leading variables from lower equations because the system is in echelon form, so leading variables in higher equation do not appear in this equation. The result 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. Move the free variables to the right side and divide by $a_{m-(t+1),\ell_{m-(t+1)}}$ to end with this equation's leading variable $x_{\ell_{m-(t+1)}}$ in terms of free variables.

We have done both the base step and the inductive step so by the principle of mathematical induction the proposition is true. $\square$

Lemma 1.9.2. *For a linear system, where $\vec{p}$ is any particular solution, the solution set equals this set.*

$$\{\vec{p}+\vec{h} \mid \vec{h} \textit{ satisfies the associated homogeneous system}\}$$

Proof. We will show mutual set inclusion, that any solution to the system is in the above set and that anything in the set is a solution of the system.

For set inclusion the first way, that if a vector solves the system then it is in the set described above, assume that $\vec{s}$ solves the system. Then $\vec{s}-\vec{p}$ solves the associated homogeneous system since for each equation index i,

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

where p_j and s_j are the j-th components of $\vec{p}$ and $\vec{s}$. Express $\vec{s}$ in the required $\vec{p}+\vec{h}$ form by writing $\vec{s}-\vec{p}$ as $\vec{h}$.

For set inclusion the other way, take a vector of the form $\vec{p}+\vec{h}$, where $\vec{p}$ solves the system and $\vec{h}$ solves the associated homogeneous system and note that $\vec{p}+\vec{h}$ solves the given system since for any equation index i,

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

where as earlier p_j and h_j are the j-th components of $\vec{p}$ and $\vec{h}$. $\square$

Corollary 1.9.1. *Solution sets of linear systems are either empty, have one element, or have infinitely many elements.*

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

1.10 Zero equals zero and number of solutions

Example

$$\left| \begin{array}{rrcl} x & +3y & = & 5 \\ 4x & +12y & = & k \end{array} \right. \quad \to \quad \left| \begin{array}{rrcr} x & +3y & = & 5 \\ 0x & +0y & = & k-20 \end{array} \right.$$

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?

$$\left| \begin{array}{rrrrcr} x & +3y & +0w & +0z & = & 5 \\ 4x & +12y & +0w & +0w & = & 20 \\ 0x & +0y & +w & +3z & = & 5 \\ 0x & +0y & +4w & +12w & = & 21 \end{array} \right. \quad \to \quad \left| \begin{array}{rrrrcr} x & +3y & +0w & +0z & = & 5 \\ 0x & +0y & +0w & +0w & = & 0 \\ 0x & +0y & +w & +3z & = & 5 \\ 0x & +0y & +0w & +0w & = & 1 \end{array} \right.$$

Observe that

$$\left| \begin{array}{rrrcr} x & +y & +z & = & 0 \\ & y & +z & = & 0 \end{array} \right.$$

has infinitely many solutions. So $0 = 0$ is not necessary for infinitely many solutions.

Remark in general we know that 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.