Chapter 1

Friday, January 2, 2026 8:02 am

Chapter 1. Linear Equations

§ 1.1 Fields

Story: It is often helpful to "abstractly" describe the key properties of real/complex numbers.

One such useful abstraction is called a "field" which captures "all" the salient features of such "scalars"

Def^n: Let F be a set &
suppose two operations "addition" &
"multiplication"
are given by $+ : F \times F \to F$ &
$\cdot : F \times F \to F$.

We say F is a field, if the following conditions hold
(together with the operations $+$ & $\cdot$).

1. Addition is commutative,
$$x + y = y + x \quad \forall\ x, y \in F.$$

2. Addition is associative,
$$x + (y + z) = (x + z) + z \quad \forall\ x, y, z \in F.$$

3. There is a unique element 0 (zero) in F
s.t. $x + 0 = x \quad \forall\ x \in F$

4. For all $x \in F$, $\exists$ a unique element $(-x) \in F$
s.t. $x + (-x) = 0$.

5. Multiplication is commutative, i.e. $xy = yx\ \forall x, y \in F$

6. Multiplication is associative, i.e.

$$x(yz) = (xy)z \quad \forall\ x, y, z \in F.$$

7. $\exists$ a unique non-zero element 1 (one) in F s.t.

$$x1 = x \quad \forall\ x \in F.$$

8. To each non-zero x in F,

$\exists$ a corresponding unique element x^{-1} (or $1/x$) in F s.t. $xx^{-1} = 1$.

9. Multiplication distributes over addition

i.e. $x(y+z) = xy + xz \quad \forall\ x, y, z \in F.$

Story: • Intuitively then, a field is a set together with some operations on this set that behave like ordinary addition (subtraction) & multiplication (division).

• We call these numbers "scalars".

NB. The set of real numbers, with usual addition & multiplication constitute a field.

Similarly for complex numbers.

Story: We now define a "subfield" which as the name suggests, is a "smaller field" contained in a given "field".

Defⁿ: Let $(F, +, \cdot)$ be a field.

Then $(F', +, \cdot)$ is a subfield of $(F, +, \cdot)$

if $F' \subseteq F$ & $(F', +, \cdot)$ is also a field.

NB. The set of real numbers is a subfield of the set of complex numbers.

(Notation: Sometimes we refer to F as the field instead of $(F, +, \cdot)$ which is the formally correct defⁿ.)

Story: The usefulness of a subfield arises from the following:

if one is working with a subfield of say $\mathbb{C}$ (complex numbers)

then performing additions (subtractions), multiplications (additions) does not take one out of the subfield.

e.g. 1. The set of positive integers:
$1, 2, 3, \ldots$ is not a subfield of $\mathbb{C}$.

(This is for a variety of reasons — 0 not in the set, missing additive & multiplicative inverses.)

e.g. 2. The set of integers: $\ldots, -2, -1, 0, 1, 2, \ldots$
is not a subfield of $\mathbb{C}$.
(This is because multiplicative inverses are missing.)
(but it satisfies the other requirements)

e.g. 3. The set of rational numbers
(i.e. numbers of the form p/q for p & q integers & $q \neq 0$)
is a subfield of $\mathbb{C}$.

NB. Every subfield of $\mathbb{C}$ must contain the set of rational numbers.

e.g. 4. The set of all complex numbers of the form
$x + y\sqrt{2}$ (where x & y are rational)
is a subfield of $\mathbb{C}$.

Assumption: Henceforth, every field considered is a subfield of the complex numbers.
(unless expressly stated otherwise)

Story: Why?
(It so happens that one can add 1 n-many times to obtain a zero.
(Will see this in Exercise 5, following §1.2)
for a finite n.
called Fields of characteristic n.

If this does not happen (as is the case for $\mathbb{C}$ & its subfields)
then F is said to have (somewhat confusingly) characteristic 0.

We won't dwell much on this here, in this course.

§ 1.2 Systems of Linear Equations

Def^n: Suppose F is a field.
A system of m linear equations in n unknowns
is any problem of the following form:
find n scalars (elements of F) $x_1 \dots x_n$

s.t.

$$(1\text{-}1)\quad \begin{aligned} A_{11}x_1 + A_{12}x_2 + \dots + A_{1n}x_n &= y_1 \\ \vdots \qquad \vdots \qquad & \quad \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \dots + A_{mn}x_n &= y_m \end{aligned}$$

where $y_1 \dots y_m$ &
$\{A_{ij}\}_{\substack{i \in \{1,\dots m\} \\ j \in \{1,\dots n\}}}$ are given elements of F.

Any n-tuple $(x_1 \dots x_n)$ of elements in F satisfying
each equation, is called a solution of the system.

If $y_1 = y_2 = \dots = y_m = 0$, the system is called
homogeneous.

Story: Can one say, given two systems of equations,
that their solutions will be the same?

We begin with an illustration & then
try to generalise the idea to
answer the question above.

Illustration: Suppose one wants to solve

$$2x_1 - x_2 + x_3 = 0$$
$$x_1 + 3x_2 + 4x_3 = 0$$

• Lets add (-7) times the second eq^n to the first to get

$$-7x_2 - 7x_3 = 0 \iff x_2 = -x_3.$$

• Lets add (3) times the first eq^n to the second to get

$$7x_1 + 7x_3 = 0 \iff x_1 = -x_3.$$

We conclude that any solution x_1, x_2, x_3 of the system must satisfy $x_1 = x_2 = -x_3$. (Conversely) One can verify that for all a, $(a, a, -a)$ is a solution.

Story (resumed): We followed the process of "eliminating unknowns".

We multiplied & added equations to remove unknowns from our equations.

Let us formalise this process slightly.

Def^n: Recall the previous def^n & (1-1) in particular. Suppose one selects m scalars $c_1 \dots c_m$, multiplies the j^{th} equation by c_j & then adds these to obtain

$$(c_1 A_{11} + \dots + c_m A_{m1}) x_1 + \dots + (c_1 A_{1n} + \dots + c_m A_{mn}) x_n$$
$$= c_1 y_1 + \dots + c_m y_m.$$

Such an equation is called a linear combination of equations in (1-1).

NB: Any solution to (1-1) is also a solution to the equation above.

Story: One can extend this idea to define a notion of equivalence of systems of linear equations.

Def^n: Recall (1-1) & consider another system of equations

$$B_{11} x_1 + \dots + B_{1n} x_n = z_1$$

$$\vdots \qquad\qquad \vdots$$

$$B_{k1}x_1 + \cdots + B_{kn}x_n = z_k.$$

These two systems are equivalent, if

each equation in each system
is a linear combination of
the equations in the other system.

Story: Clearly, if the two systems are equivalent
& $x_1 \dots x_n$ is a solution to the first system
it is also a solution to each equation in the second system.

Conversely, if $x_1 \dots x_n$ is a sol^n to the second
it is also a sol^n to the first.

We therefore have the following.

Theorem 1. Equivalent systems of linear equations have exactly the same solutions.

Story: In the next section, we will see how to produce an equivalent system that is easier to solve, starting from an arbitrary initial system.

<u>Rough</u>

$x + y\sqrt{2}$ addition / subtraction / multi^n

Division ~~suppose~~ Fix x, y, find p, q

$$(x + y\sqrt{2})(p + q\sqrt{2}) = 1$$

$$xp + 2yq + (\cancel{x}py + xq\cancel{q})\sqrt{2} = 1$$

$$xp + 2yq = 1$$

$$\cancel{xp}\,(py + xq) = 0$$

$$xp + 2yq = 1$$

$$p = -\frac{xq}{y}$$

$$-x\left(\frac{xq}{y}\right) + 2yq = 1$$

$$\left(-\frac{x^2}{y} + 2y\right)q = 1$$

$$q = \frac{1}{2y - \frac{x^2}{y}}$$

$$p = -\frac{xq}{y}$$

§1.3 Matrices and Elementary Row Operations

Story: We now make our notation more brief.

Notation: We abbreviate the system of equations (1-1) by

$$AX = Y$$

where $A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$ $\quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ $\quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$

A is called the matrix of coefficients of the system

Remarks:

(1) For now, we take this to simply be a shorthand

We will look at the matrix multiplication notation and definition later.

(2) A is strictly speaking, not a matrix

An $m \times n$ matrix over the field F

is a function A from (i, j) where $1 \leq i \leq m$ and $1 \leq j \leq n$

into the field F

Entries of the matrix A are the scalars $A(i,j) = A_{ij}$

Story:

We now want to consider operations
on the rows of the matrix A
that correspond to forming linear combinations
of the equations in the system
$AX = Y$

We first look at three elementary row operation.

Inf Definition. **Elementary row operations** on an $m \times n$ matrix A over the field F are defined to be the following.

1. Multiplication of one row of A by a non-zero scalar c
2. Replacement of the rth row of A
   by row r plus c times row s
   where c is any scalar and $r \neq s$
3. Interchange of two rows of A

Story:

An elementary row operation is
thus a special type of function (or rule) e
that associates with each $m \times n$ matrix A
another $m \times n$ matrix $e(A)$.

Using this, one can define these operations more formally.

Definition. Elementary row operations.

Let A be an $m \times n$ matrix as above and
let e be an arbitrary function of the form above.

Then, e must satisfy one of the following
(for some row distinct row indeces r, s and scalar $c \neq 0$):

1. $e(A)_{ij} = A_{ij}$ for $i \in \{1 \dots m\} \backslash r$ and $e(A)_{rj} = cA_{rj}$
2. $e(A)_{ij} = A_{ij}$ for $i \in \{1 \dots m\} \backslash r$ and $e(A)_{rj} = A_{rj} + cA_{sj}$
3. $e(A)_{ij} = A_{ij}$ for $i \in \{1 \dots m\} \backslash \{r, s\}$ and $e(A)_{rj} = A_{sj}$, $e(A)_{sj} = A_{rj}$

These are called types of "e"

Story:

In defining $e(A)$
it is not really important how many columns A has (i.e. n can be arbitrary)

but th enumber of rows of A (i.e. m) is crucial
(in the sense that interchanging rows 5 and 6
of a 5 x 6 matrix makes no sense).

Therefore when we speak of e, we are considering the class of all m-rowed matrices over F.

Why these specific operations?
One reason is that after performing e on a matrix A to obtain $e(A)$,
one can recover A by performing a similar operation on $e(A)$.

Theorem 2.
To each elementary row operation e
there corresponds an elementary row operation e_1
of the same type as e
such that $e_1\big(e(A)\big) = e\big(e_1(A)\big)$ for all A.

In other words, the inverse operation (function)
of an elementray row operation exists
and it is an elementary row operation
of the same type.

Proof.

We do the proof for each type separately

(1) Suppose e is the operation which
multiplies the rth row of a matrix
by the non-zero scalar c.

Clearly, we can take e_1 to be the operation that multiplies r by c^{-1}.

(2) Suppose e is the operation which replaces row r by row r plus c times row s
($s \neq r$).
Again, clearly, we can take e_1 to be the operation that
replaces row r by row r plus $(-c)$ times row s.

(3) If e interchanges rows r and s, we can take $e_1 = e$.

It is easy to check that we have $e_1\big(e(A)\big) = e\big(e_1(A)\big) = A$ for all A.

■

Story: We now use these operations to define another notion of equivalence between matrices.

Definition. If A and B are $m \times n$ matrices over the field F
we say that B is row-equivalent to A

if B can be obtained from A
by a finite sequence of elementary row operations.

Story: Here are some simple observations that can be verified using Theorem 2.

NB.

(1) Each matrix is row-equivalent to itself

(2) if B is row-equivalent to A then A is row-equivalent to B

(3) if C is row-equivalent to B and B is row-equivalent to A then C is row-equivalent to A.

In other words, row-equivalence is an equivalence relation.

Theorem 3. If A and B are row-equivalent $m \times n$ matrices the homogenous systems of linear equaions $AX = 0$ and $BX = 0$ have exactly the same solutions.

Proof.

Suppose we pass from A to B by a finite sequence of elementray row operations:

$$A = A_0 \to A_1 \to \cdots \to A_k = B$$

NB. It is enough to prove that the systems $A_j X = 0$ and $A_{j+1} X = 0$ have the same solutions, i.e. one elementary row operation leaves the set of solutions unchanged.

This is easy to establish.

Suppose C is obtained from D by a single elementary operation (where C and D are also $m \times n$ matrices)

Now, no matter which of the three types of elementary row operation is used it is clear that each equation in system $CX = 0$ is a linear combination of th eequatinos in $AX = 0$

Furthermore, since the inverse of an elementary row operation is also an elementary row operation each equation in $AX = 0$ will also be a linear combination of the equations in $BX = 0$.

Thus, these two systems are equivalent—and by Theorem 1, have the same solutions.

■

Story:

Now let us look at how one might use these elementary row-operations to potentially solve a homogenous system of linear equations

EXAMPLE 5. Suppose F is the field of rational numbers, and

$$A = \begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix}.$$

We shall perform a finite sequence of elementary row operations on A, indicating by numbers in parentheses the type of operation performed.

$$\begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & -2 & -1 & 7 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 0 & \frac{15}{2} & -\frac{55}{2} \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(1)}$$

$$\begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{5}{3} \end{bmatrix}$$

The row-equivalence of A with the final matrix in the above sequence tells us in particular that the solutions of

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

and

$$\begin{aligned} x_3 - \tfrac{11}{3}x_4 &= 0 \\ x_1 + \tfrac{17}{3}x_4 &= 0 \\ x_2 - \tfrac{5}{3}x_4 &= 0 \end{aligned}$$

are exactly the same. In the second system it is apparent that if we assign

any rational value c to x_4 we obtain a solution $(-\frac{17}{3}c, \frac{5}{3}, \frac{11}{3}c, c)$, and also that every solution is of this form.

Here's another example.

EXAMPLE 6. Suppose F is the field of complex numbers and

$$A = \begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix}.$$

In performing row operations it is often convenient to combine several operations of type (2). With this in mind

$$\begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 2+i \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & 1 \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Thus the system of equations

$$\begin{aligned} -x_1 + ix_2 &= 0 \\ -ix_1 + 3x_2 &= 0 \\ x_1 + 2x_2 &= 0 \end{aligned}$$

has only the trivial solution $x_1 = x_2 = 0$.

Story:
In these examples,
we were trying to simplify the coefficient matrix
in a manner analogous to "eliminating unkowns"
in the system of linear equations.

Let us now make a formal definition of
the type of matrix at which
we were attempting to arrive.

Definition. An $m \times n$ matrix R is called **row-reduced** if
(1) the first non-zero entry in
each non-zero row of R is 1
(2) each column of R
which contains th leading non-zero entry of some row
has all its other entries 0.

Story: Let us look at a very simple example of a row-reduced matrix.

EXAMPLE 7. One example of a row-reduced matrix is the $n \times n$ (square) **identity matrix** I. This is the $n \times n$ matrix defined by

$$I_{ij} = \delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j. \end{cases}$$

This is the first of many occasions on which we shall use the **Kronecker delta** (δ).

Story:

Note that in Examples 5 and 6,
the final matrices in the sequence
were row-reduced matrices.

$$\begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{5}{3} \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Here are some examples of matrices
that are *not* row-reduced.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix}.$$

The first matrix fails condition (b) in column 3
The second matrix fails condition (a) in row 1

Theorem 4.

Every $m \times n$ matrix over the field F is
row-equivalent to
a row-reduced matrix.

Proof.

Let A be an $m \times n$ matrix over F.

Handling row 1
If every entry in the first row of A is 0
then condition (a) is satisfied
as far as row 1 is concerned.
If row 1 has a non-zero entry
let k be the smallest index for which $A_{ik} \neq 0$

Multiply row 1 by A_{1k}^{-1}
and then condition (a) is satisfied with regard to row 1.

Now, for each $i \geq 2$
add $(-A_{ik})$ times row 1 to row i.

At this point, the leading non-zero entry of row 1
occurs in column k and that entry is 1.
Also, all other entries in column k are 0.

Handling row 2
Consider now the matrix which we obtained above.

If every entry in row 2 is 0, we leave it unchanged.
If some entry in row 2 is not zero
we multiply row 2 by a scalar so that the leading non-zero term
is again 1.

Suppose this leading entry occurs at row k_r

It is clear that $k' \neq k$ (from how we handled row 1; recall the definiton of k above)

By adding suitable multiples of row 2 to the various rows
one can ensure that all entries in column k'
are 0 except for row 2 (where it takes value 1).

The crucial observation is that
in carrying out these operations,
the entries of row 1 will remain unchanged
and
the entries of column k remain unchanged.

Of course, if row 1 is identically 0, the operations with row 2
will not affect row 1.

Proceeding like this, one row at a time, it is clear that
in a finite number of steps, one will arrive at a row-reduced matrix.