

Linear Algebra Notes — Lecture 8/25

Linear Systems, Row Operations and Reduced Echelon Form

Lecture pages 1–8 and 12 · Lay §1.1–1.2

Contents

Overview	1
1 Linear equations (pp. 1–2)	2
2 Systems of linear equations (p. 3)	3
3 A heuristic in $\mathbb{R}^2$: zero, one or infinitely many solutions (pp. 4–5)	5
4 The augmented matrix (p. 5)	6
5 Row operations do not change the solutions (p. 6)	6
6 Reduced row echelon form (pp. 7, 12)	8
7 The row reduction algorithm (pp. 7–8)	9
Summary	10
Corrections to the handwritten notes	11

Overview

This is the first lecture of the course. It sets up the central problem of the first weeks: *solve a system of linear equations*. The plan is:

1. Define linear equations and systems, and picture them as lines and planes (pp. 1–4).
2. Notice that only the numbers matter, and store them in the *augmented matrix* $[A \mid b]$ (p. 5).
3. Simplify $[A \mid b]$ with three *row operations*, which never change the solutions (p. 6).
4. Aim for one particular simple shape, the *reduced row echelon form* (RREF), and give an algorithm that always gets there (pp. 7–8, 12).

Every blank and question in the notes is answered in a blue box. Typos are listed at the end.

Notation used throughout.

- a_{ij} : the coefficient in equation i of the variable x_j (row i , column j).
- $[A \mid b]$: the augmented matrix of a system; A is the coefficient matrix, b the right-hand side.
- $Ax = b$: shorthand for the whole system (p. 5).
- RREF: reduced row echelon form (p. 7). A *pivot* is a leading 1 of an RREF matrix.
- $\mathbb{R}^n$: the set of columns $(x_1, \dots, x_n)$ of n real numbers (formally defined on p. 14 of the 8/27 notes).

Recipe of the lecture. To solve a linear system: write $[A \mid b]$; row reduce it to RREF $[R \mid d]$ (algorithm on pp. 7–8); then solve $Rx = d$, which is easy and has the *same* solutions as $Ax = b$ (Theorem, p. 6). Reading off the solutions of $Rx = d$ is done in the 8/27 lecture (p. 9).

1 Linear equations (pp. 1–2)

A linear equation is a sum of constants times variables set equal to a constant; over the real numbers we can draw its solutions.

Definition (linear equation, p. 1)

A **linear equation** in the variables (unknowns) $x_1, \dots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

where the numbers $a_1, \dots, a_n$ are the **coefficients** and b is the **right-hand side**.

“Linear” means each variable appears only to the first power, multiplied by a constant: no x_1^2 , no x_1x_2 , no $\sin x_1$. For example $3x_1 - x_2 = 7$ is linear; $x_1x_2 = 1$ is not.

Which numbers? For now all a_i, x_i, b are real numbers. But the definition only uses addition and multiplication, so it makes sense for other number systems (“fields”) too:

- the complex numbers $\mathbb{C}$;
- finite fields, for example $\mathbb{Z}/2\mathbb{Z} = \{0, 1\}$ (“the integers mod 2”) with the rules

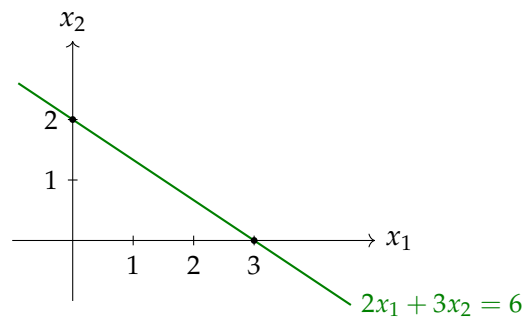
$$\begin{aligned} 0 + 0 &= 0, & 1 + 0 &= 0 + 1 = 1, & 1 + 1 &= 0, \\ 0 \cdot 0 &= 0, & 0 \cdot 1 &= 1 \cdot 0 = 0, & 1 \cdot 1 &= 1. \end{aligned}$$

The only surprising rule is $1 + 1 = 0$: think of 0 as “even” and 1 as “odd” (odd + odd = even). *Example (an addition).* Over $\mathbb{Z}/2\mathbb{Z}$ the equation $x_1 + x_2 = 1$ has exactly two solutions, $(x_1, x_2) = (1, 0)$ and $(0, 1)$: the other two candidates give $0 + 0 = 0$ and $1 + 1 = 0$. So over a finite field a linear equation has finitely many solutions, unlike over $\mathbb{R}$.

Pictures over the real numbers (p. 2)

Real numbers let us *visualize* solutions.

Example: $2x_1 + 3x_2 = 6$ is a line in $\mathbb{R}^2$. Solving for x_2 gives $x_2 = 2 - \frac{2}{3}x_1$: slope $-\frac{2}{3}$, crossing the x_2 -axis at 2 (set $x_1 = 0$) and the x_1 -axis at 3 (set $x_2 = 0$).



The green line is the set of all solutions; it passes through $(0,2)$ and $(3,0)$. Check: $2 \cdot 0 + 3 \cdot 2 = 6$ and $2 \cdot 3 + 3 \cdot 0 = 6$. ✓

In many applications we do not need the picture. The algebra keeps working even when the geometry is hard to imagine (e.g. 100 variables).

Example: $2x_1 + 3x_2 + 0x_3 = 6$ is a plane in $\mathbb{R}^3$ with normal vector $\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$. It passes through the points

$$\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 6 \\ -2 \\ 0 \end{bmatrix}.$$

Check: $2 \cdot 3 + 0 = 6$, $3 \cdot 2 = 6$, $2 \cdot 6 + 3 \cdot (-2) = 12 - 6 = 6$. ✓

Where the normal vector comes from (from multivariable calculus, not from the notes). The notes rewrite the equation as a “row times column”:

$$\begin{bmatrix} 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 2x_1 + 3x_2 + 0x_3 = 6.$$

The left side is the dot product $n \cdot x$ with $n = (2, 3, 0)$, the coefficients of the equation. If p and q are two points of the plane, then $n \cdot (p - q) = 6 - 6 = 0$, so n is perpendicular to every direction inside the plane: that is what “normal vector” means. Since the coefficient of x_3 is 0, x_3 can be anything, so the plane is “vertical”: it is the line of the previous example, swept up and down the x_3 -axis.

2 Systems of linear equations (p. 3)

A system is several linear equations that must hold at the same time.

Example from the notes:

$$\begin{cases} 2x_1 + 3x_2 + 0x_3 = 6 \\ -x_1 + 2x_2 - x_3 = 2 \end{cases}$$

Geometrically, each equation is a plane in $\mathbb{R}^3$, and the solutions are the points on both planes. The normals $(2, 3, 0)$ and $(-1, 2, -1)$ are not parallel, so the planes are not parallel and meet in a line.

Definition (system of linear equations, p. 3)

A system of m linear equations in the unknowns $x_1, \dots, x_n$ is

$$(*) \quad \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m. \end{cases}$$

A **solution** is a list $(x_1, \dots, x_n)$ of numbers satisfying all m equations at once.

The double index in a_{ij} records position: i is the equation (row), j the variable (column).

Questions (the notes' two guiding questions for the next lectures).

Question 1: How do we describe the solutions of $(*)$? By row reducing to RREF

This lecture gives the method: replace the system by a simpler one with the same solutions (Theorem, p. 6), namely its RREF (pp. 7–8). In RREF each equation contains one “pivot” variable that can be solved for directly. The 8/27 notes (p. 9) then write all solutions in *parametric form*. A preview from p. 4: the system $x_1 + x_2 = 1$, $2x_1 + 2x_2 = 2$ has the solutions $\begin{bmatrix} x_1 \\ 1-x_1 \end{bmatrix}$, one for each value of x_1 .

Question 2: How many solutions can $(*)$ have? 0, 1 or infinitely many

The pictures on p. 4 show all three cases for two equations in two unknowns: two lines meet in one point, are identical, or are parallel. Nothing else can happen in $\mathbb{R}^2$: two distinct lines meet in at most one point. The same three possibilities hold for every linear system, and the RREF decides which case occurs. This is proved in the 8/27 lecture (pp. 10–11; Lay §1.2, Theorem 2). *Quick reason it can't be “exactly 2” (an addition):* if $p \neq q$ are two solutions of $(*)$, then for every real t the point $p + t(q - p)$ is also a solution, because each equation gives $a_{i1}(p_1 + t(q_1 - p_1)) + \dots = b_i + t(b_i - b_i) = b_i$. So two solutions force infinitely many.

Why do we care? Every system is locally approximately linear

The notes put it in one line: “Using multivariable Taylor expansion any system of equations can be approximated locally by a linear system.” Near a point a , a smooth function is approximately its first-order Taylor polynomial, $f(x) \approx f(a) + \nabla f(a) \cdot (x - a)$, which is linear in x . *Example (an addition):* near $(1, 1)$, the circle $x^2 + y^2 = 2$ is approximated by $2 + 2(x - 1) + 2(y - 1) = 2$, i.e. the linear equation $x + y = 2$: the tangent line at $(1, 1)$. *Check:* $(1, 1)$ satisfies both, and the point $(1.01, 0.99)$ on the line gives $x^2 + y^2 = 2.0002 \approx 2$. ✓ So understanding linear systems is the first step to understanding all systems of equations.

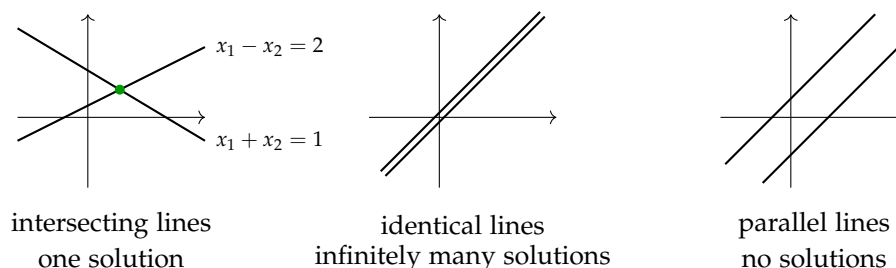
3 A heuristic in $\mathbb{R}^2$: zero, one or infinitely many solutions (pp. 4–5)

With $m = n = 2$, each equation is a line, and the solutions of the system are the points on both lines.

The solutions of one equation $a_{i1}x_1 + a_{i2}x_2 = b_i$ form a line in $\mathbb{R}^2$, or the empty set. *Example:* $0 \cdot x_1 + 0 \cdot x_2 = 1$ has no solutions, since the left side is always 0.

Fine print. There is a third case: if $a_{i1} = a_{i2} = b_i = 0$, the equation $0x_1 + 0x_2 = 0$ holds for *every* point, so its solution set is all of $\mathbb{R}^2$. If at least one of a_{i1}, a_{i2} is nonzero, the solution set is a line.

Ignoring the empty set for now, a solution of a system of two equations in two variables is a point $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ that lies on both lines. There are three pictures:



The three sketches are schematic (as in the notes): two lines crossing in one green point, two lines lying on top of each other (drawn as one doubled line), and two parallel lines.

Picture	Example from the notes	Solutions
intersecting lines	$x_1 + x_2 = 1, x_1 - x_2 = 2$	one: $(x_1, x_2) = (\frac{3}{2}, -\frac{1}{2})$
identical lines	$x_1 + x_2 = 1, 2x_1 + 2x_2 = 2$	infinitely many: $\begin{bmatrix} x_1 \\ 1-x_1 \end{bmatrix}, x_1 \text{ any number}$
parallel lines	$x_1 + x_2 = 1, x_1 + x_2 = 0$	none

Identical lines. The second equation is $2\times$ the first, so it says nothing new. Every solution of $x_1 + x_2 = 1$ works: pick x_1 freely and set $x_2 = 1 - x_1$. *Check:* $x_1 + (1 - x_1) = 1$ and $2x_1 + 2(1 - x_1) = 2$. ✓

Parallel lines. $x_1 + x_2$ cannot equal 1 and 0 at the same time, so there is no solution.

Intersecting lines (solved on p. 5). Subtract equation 1 from equation 2: $(x_1 - x_2) - (x_1 + x_2) = 2 - 1$, i.e. $-2x_2 = 1$. The system becomes

$$\begin{cases} x_1 + x_2 = 1 \\ -2x_2 = 1 \end{cases} \implies x_2 = -\frac{1}{2}, \quad x_1 = 1 - x_2 = \frac{3}{2}.$$

The arrow in the notes shows the order: solve the bottom equation for x_2 , then substitute it into the top one (“back substitution”). *Check:* $\frac{3}{2} + (-\frac{1}{2}) = 1$ and $\frac{3}{2} - (-\frac{1}{2}) = 2$. ✓

4 The augmented matrix (p. 5)

Only the numbers a_{ij} , b_i matter, so we store them in a rectangular table.

Goal. Develop an algorithm for solving linear systems that can be made into a computer program.

In $(*)$, only the numbers a_{ij} and b_i matter; the indices i, j and i just encode their positions. So we write them in a grid:

Definition (augmented matrix, coefficient matrix, p. 5)

The **augmented matrix** of $(*)$ is

$$[A | b] = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right].$$

Row i holds equation i ; column j (for $j \leq n$) holds the coefficients of x_j ; the last column holds the right-hand sides. Here A (the $m \times n$ block on the left) is the **matrix of coefficients** and b (the last column) is the **right-hand side** of $(*)$.

Example. The intersecting-lines system above has $[A | b] = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & -1 & 2 \end{array} \right]$; the system on p. 3 has $[A | b] = \left[\begin{array}{ccc|c} 2 & 3 & 0 & 6 \\ -1 & 2 & -1 & 2 \end{array} \right]$ (a 0 is written for the missing x_3 in the first equation).

Strategy. Do row operations (defined on p. 6) on $[A | b]$ to bring it to a simpler matrix.

Notation $Ax = b$. We write the system $(*)$ as $Ax = b$, where $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

“The notation will make sense later.” For now $Ax = b$ is just an abbreviation of $(*)$. Soon (8/27 notes, p. 18; Lay §1.4) Ax will be defined as a product of the matrix A and the column x , and then $Ax = b$ will literally be an equation between columns. The “row times column” on p. 2, $\begin{bmatrix} 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 2x_1 + 3x_2 + 0x_3$, is the one-row case: row i of A times x is the left side of equation i .

5 Row operations do not change the solutions (p. 6)

Each row operation can be undone, so it produces a system with exactly the same solutions.

Definition (elementary row operations, p. 6)

1. **Replacement:** add a multiple of one row to another row.
2. **Interchange:** swap two rows.
3. **Scaling:** multiply a row by a *nonzero* number.

On the system, these are: add c times equation j to equation i ($j \neq i$); list the equations in a different order; multiply both sides of an equation by $c \neq 0$.

Key feature: each row operation is reversible, so no information is lost:

- adding $c \cdot (\text{row } j)$ to row i is undone by adding $-c \cdot (\text{row } j)$ to row i (row j itself did not change);
- swapping rows i and j is undone by swapping them again;
- multiplying row i by $c \neq 0$ is undone by multiplying it by $1/c$. (This is why $c = 0$ is not allowed: multiplying by 0 erases the equation and cannot be undone.)

Theorem (p. 6)

If $[R \mid d]$ is obtained from $[A \mid b]$ by a sequence of row operations, then

$$Ax = b \iff Rx = d,$$

that is, the systems $Ax = b$ and $Rx = d$ have the same solutions.

Proof (the “clearly” and Operation 1 filled in)

Induction on $r =$ the number of row operations.

Base case $r = 0$: no operations, so $[R \mid d] = [A \mid b]$ and there is nothing to prove.

Induction step $r \rightarrow r + 1$: the first r operations give a system with the same solutions as $Ax = b$ (induction hypothesis), so it is enough to show that a *single* row operation does not change the solutions.

- *Operation 2 (interchange)*: the same equations are listed in a different order, and “all equations hold” does not depend on the order.
- *Operation 3 (scaling by $c \neq 0$)*: if $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$, then multiplying by c gives $ca_{i1}x_1 + \cdots + ca_{in}x_n = cb_i$. Conversely, multiplying the new equation by $1/c$ gives back the old one. So x satisfies the old equation exactly when it satisfies the new one.
- *Operation 1 (replacement)*: the notes say this is “simply adding b_j to both sides of $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ ”. In detail, adding c times row j to row i changes equation i to

$$(a_{i1} + ca_{j1})x_1 + \cdots + (a_{in} + ca_{jn})x_n = b_i + cb_j.$$

If x solves the old system, then $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$ (equation j), so the left side gains $c(a_{j1}x_1 + \cdots + a_{jn}x_n) = cb_j$, the same amount as the right side: the new equation holds. All other equations are unchanged. Conversely, the old system is obtained from the new one by adding $-c$ times row j (which is still in the new system), so the same argument shows every solution of the new system solves the old one.

Hence after $r + 1$ operations the solutions are still those of $Ax = b$. □

Example. On p. 5 we subtracted equation 1 from 2 (replacement with $c = -1$) and the solution $(\frac{3}{2}, -\frac{1}{2})$ of the original system also solves the new one: $-2 \cdot (-\frac{1}{2}) = 1$. ✓

6 Reduced row echelon form (pp. 7, 12)

We row reduce towards one specific simple shape, the RREF, from which solutions can be read off.

What RREF looks like. The notes give this example, where each * stands for any number:

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

Reading it: the leading 1s (in columns 2,4,6,7) move strictly to the right as we go down, like a staircase; every other entry in a column containing a leading 1 is 0; the zero row is at the bottom. The * entries sit only in columns without a leading 1 (or to the right of the leading 1 in its row).

Definition (RREF, p. 7)

A matrix is in **reduced row echelon form** (RREF) if

1. all zero rows (if any) are at the bottom;
2. the leftmost nonzero entry in each nonzero row equals 1 (a **pivot**), and it is to the right of the pivots in the rows above (if any);
3. the pivots are the only nonzero entries in their columns.

Definition (pivotal, p. 7)

The row, column and position of a pivot are called a **pivotal** row, column and position (Lay: pivot position, pivot column).

In the example above the pivotal columns are 2,4,6,7 and the pivotal rows are 1–4.

Examples (p. 12): is it in RREF?

Is it in RREF? $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$: No; $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$: No; $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$: Yes

The notes give the answers; here are the reasons.

- $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$: **No**. The pivot of row 2 is in column 2, and the entry above it is $1 \neq 0$, so rule 3 fails. (One replacement, row 1 – row 2, gives the RREF $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.)
- $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$: **No**. The zero row is at the top, so rule 1 fails. (Moving it to the bottom with two swaps gives the RREF $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.)
- $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$: **Yes**. The zero row is at the bottom, the only pivot is the 1 in position (1,1), and column 1 has no other nonzero entry. The 2 is allowed: column 2 is not a pivot column.

Check (by computer, *sympy's rref*): the RREFs of the three matrices are $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and the third matrix itself. ✓

About page 12. This page carries the professor's page number 12 but is filed with the 8/25 notes; the 8/27 notes jump from p. 11 to p. 13. It is covered here, after the definition of RREF that it tests.

7 The row reduction algorithm (pp. 7–8)

A fixed sequence of row operations turns any matrix into RREF in finitely many steps.

Row reduction algorithm (pp. 7–8)

1. Cover the zero columns on the left.
2. Look at the leftmost nonzero column. Pick a nonzero entry in that column.
 - 2a Move it to the top via a row swap.
 - 2b Scale it to be 1 (this is the new pivot).
 - 2c Use row replacement to make all other entries in the column = 0.
3. Cover the column and the top row to get a smaller matrix, and repeat the process.

Reading step 2c. To end in RREF (rule 3), the entries *above* each new pivot must also be cleared, i.e. step 2c works on the whole column of the full matrix, including the covered rows. The example below does exactly this in its last step. Lay (§1.2) instead first clears only below the pivots ("forward phase", giving an echelon form) and then clears above them from right to left ("backward phase"); both give the same result.

Why does the row reduction algorithm terminate? Each pass covers a row

Each pass through steps 1–3 does finitely many row operations: at most one swap, one scaling, and at most $m - 1$ replacements (one per other row). At the end of the pass, step 3 covers one row. A matrix with m rows can lose a row only m times, and the process also stops earlier if the uncovered part has no nonzero column left. So there are at most m passes and at most $m(m + 1)$ row operations in total. When it stops, every covered row has a pivot 1 that is the only nonzero entry in its column, the pivots move right as we go down (each new pivot column lies to the right of the covered columns), and the remaining rows are zero: the matrix is in RREF.

Example (p. 8): row reduce $\begin{bmatrix} 2 & 4 & 6 \\ -1 & -2 & -1 \end{bmatrix}$ in two ways

First way (pick the entry -1 in step 2). Multiply row 2 by -1 and swap the rows, so that the chosen entry becomes a pivot 1 at the top:

$$\begin{aligned} \begin{bmatrix} 2 & 4 & 6 \\ -1 & -2 & -1 \end{bmatrix} &\xrightarrow{R_2 \times (-1), R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 6 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix} \\ &\xrightarrow{R_2 \times \frac{1}{4}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 - R_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

After the first pass, column 1 is done and row 1 is covered; the smaller matrix is the row $[0 \ 0 \ 4]$, whose leftmost nonzero column is column 3 (column 2 is a zero column there and gets covered in step 1). The last step clears the 1 above the new pivot.

Second way (pick the entry 2 in step 2):

$$\begin{aligned} \begin{bmatrix} 2 & 4 & 6 \\ -1 & -2 & -1 \end{bmatrix} &\xrightarrow{R_1 \times \frac{1}{2}} \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -1 \end{bmatrix} \xrightarrow{R_2 + R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix} \\ &\xrightarrow{R_2 \times \frac{1}{2} \text{ (scale row 2)}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 - 3R_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Check: both ways end at the same matrix $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, which sympy's `rref` also returns (pivot columns 1 and 3). It satisfies rules 1–3. ✓

Meaning: the choices made in step 2 change the route but not the destination. This is a theorem: every matrix has exactly one RREF (Lay §1.2, Theorem 1; not proved in these notes).

Summary

Fact	Formula / statement
Linear equation	$a_1x_1 + \cdots + a_nx_n = b$
Number of solutions of a system	0, 1 or infinitely many (pictures in $\mathbb{R}^2$, p. 4)
Augmented matrix	$[A \mid b]$: coefficients a_{ij} in row i , column j ; right sides in the last column
Row operations	replacement, interchange, scaling by $c \neq 0$; all reversible
Theorem (p. 6)	row operations take $[A \mid b]$ to $[R \mid d] \Rightarrow Ax = b$ and $Rx = d$ have the same solutions
RREF	zero rows at bottom; leading entries are 1s moving right; pivots alone in their columns
Row reduction	terminates after at most m passes; the RREF does not depend on the choices made

Corrections to the handwritten notes

Page	As written	Should be	Why
4	solutions of one equation: “a line in $\mathbb{R}^2$ or the empty set”	a line, the empty set, or all of $\mathbb{R}^2$	$0x_1 + 0x_2 = 0$ holds for every point
4	points “that line on two lines”	that lie on two lines	word slip
6	Operation 1 is “adding b_j to both sides”	adding $c(a_{j1}x_1 + \dots + a_{jn}x_n)$ to the left and cb_j to the right	the two amounts are equal only for solutions of equation j ; the multiple c is general