

Linear Algebra Notes — Lecture 8/27

Solution Sets, Vectors, Span, and the Equation $Ax = b$

Lecture pages 9–11, 13–21 · Lay §1.2–1.4

Contents

Overview	1
1 Solving $Rx = d$: the parametric form (p. 9)	2
2 How many solutions? (pp. 10–11)	3
3 Interpolating polynomials (p. 13)	4
4 Vectors in $\mathbb{R}^n$ (pp. 14–15)	5
5 Linear combinations and span (pp. 15–17)	7
6 The matrix–vector product (p. 18)	9
7 $Ax = b$ as a vector equation; when is it always solvable? (pp. 19–21)	11
Summary	13
Corrections to the handwritten notes	13

Overview

On 8/25 we learned to row reduce an augmented matrix $[A \mid b]$ to its reduced row echelon form $[R \mid d]$, and that $Ax = b$ and $Rx = d$ have the same solutions (8/25 notes, p. 6). This lecture reads the answer off $[R \mid d]$ and then changes the point of view from equations to vectors:

1. **Reading off solutions.** Pivotal variables are solved for in terms of free variables. This gives the *parametric form* of the solution set, and a theorem saying when there are no, one, or infinitely many solutions.
2. **An application.** Fitting a polynomial through data points is a linear system.
3. **Vectors.** Vectors in $\mathbb{R}^n$, the parallelogram rule, linear combinations, and the span.
4. **$Ax = b$ as a vector equation.** The product Ax is a linear combination of the columns of A . So $Ax = b$ is solvable exactly when b is in the span of the columns, and it is solvable for

every b exactly when A has a pivot in every row.

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

Notation used throughout.

- $[A \mid b]$: the augmented matrix of the system $Ax = b$; $[R \mid d]$: its RREF (8/25 notes, pp. 5–7).
- A *pivot* is a leading 1 of a row of an RREF matrix (8/25 notes, p. 7). A *pivotal* variable or column is one whose column contains a pivot; the other variables are *free*.
- $\text{span}\{v_1, \dots, v_n\}$: the set of all linear combinations of $v_1, \dots, v_n$ (p. 15).
- $e_m = (0, \dots, 0, 1)^T$: the column vector with a single 1 at the bottom (used on p. 21).

1 Solving $Rx = d$: the parametric form (p. 9)

Solve for the pivotal variables; the free variables can be anything. The result writes every solution as one particular solution plus free-variable multiples of fixed vectors.

How to solve $Ax = b$, or equivalently $Rx = d$. “Equivalently” is the theorem from 8/25 (p. 6): row operations do not change the solution set, so we can work with the much simpler $[R \mid d]$.

Example (p. 9)

$$[R \mid d] = \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array} \right] \quad (\text{columns belong to } x_1, x_2, x_3),$$

which is the system

$$\begin{cases} x_1 + 2x_2 + 0x_3 = 3, \\ 0x_1 + 0x_2 + 1 \cdot x_3 = 2. \end{cases}$$

The pivots are in columns 1 and 3, so x_1, x_3 are the **pivotal variables** and x_2 is **free**. Each equation solves for one pivotal variable:

$$\begin{array}{lcl} x_1 = 3 - 2x_2 & & \\ x_2 = x_2 & \text{or} & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}. \\ x_3 = 2 & & \end{array}$$

The line “ $x_2 = x_2$ ” just records that x_2 is a free parameter. The vector $(3, 0, 2)$ collects the constants and $(-2, 1, 0)$ collects the coefficients of x_2 .

Meaning. The solution set is the line in $\mathbb{R}^3$ through the point $(3, 0, 2)$ in the direction of $(-2, 1, 0)$. This is the **parametric form** of the solution of $Ax = b$ (Lay §1.2 and §1.5).

Check: $R(3, 0, 2)^T = (3, 2)^T = d$, and $R(-2, 1, 0)^T = (-2 + 2, 0)^T = (0, 0)^T$, so moving along the direction vector does not change Rx . ✓

2 How many solutions? (pp. 10–11)

Everything is decided by where the pivots of $[R \mid d]$ are: a pivot in the last column means no solution; otherwise each free variable is one free parameter.

Two more examples (p. 10)

No solutions. $[R \mid d] = \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 1 \end{array} \right]$ is the system $x_1 + 2x_2 = 0$, $0 \cdot x_1 + 0 \cdot x_2 = 1$. The second equation says $0 = 1$, so there are **no solutions**.

Exactly one solution. $[R \mid d] = \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -1 \end{array} \right]$ says directly $x_1 = 2$, $x_2 = -1$. There are no free variables, so this is the only solution.

Theorem (p. 10; Lay §1.2, Theorem 2)

Let $Ax = b$ be a linear system and let $[R \mid d]$ be the RREF of $[A \mid b]$. Then:

- (1) $Ax = b$ has no solutions $\iff$ the d column contains a pivot.
- (2) If $Ax = b$ has a solution, then
 - (2a) the solution is unique $\iff$ each column of R is pivotal;
 - (2b) there are infinitely many solutions $\iff$ some column of R is not pivotal.

In particular, a linear system has 0, 1 or infinitely many solutions; never exactly 2 (compare the three pictures of two lines in $\mathbb{R}^2$, 8/25 notes, p. 4).

Proof (p. 11: “This is nearly obvious”), with the details filled in

(1) If the d column contains a pivot, that pivot is a leading 1 in some row whose entries in the R part are all 0 (a pivot is the first nonzero entry of its row). That row is the equation

$$0 \cdot x_1 + \cdots + 0 \cdot x_n = 1,$$

which no x satisfies. So there are no solutions.

If the d column contains no pivot, every nonzero row of $[R \mid d]$ has its pivot in some variable's column. Solve each such row for its pivotal variable, as on p. 9. Choosing any values for the free variables (if any) then gives a solution. So there is at least one solution. This proves both directions of (1).

(2) Suppose there is no pivot in the d column, and write the solution as on p. 9:

$$x = (\text{constant vector}) + (\text{free variable}) \cdot (\text{vector}) + \cdots$$

- If there are no free variables (every column of R is pivotal), each variable equals a fixed number, so the solution is **unique**. This is (2a).
- If some column of R is not pivotal, its variable is free. Different values of that variable give

different solutions (that variable's own entry changes), so there are **infinitely many**. This is (2b).

So the solution is unique iff there are no free variables, which is the notes' one-line summary.

Example with three free variables (p. 11)

$$[R \mid d] = \left[\begin{array}{ccccc|c} 0 & 1 & 2 & 0 & 3 & 4 \\ 0 & 0 & 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \quad (\text{columns belong to } x_1, \dots, x_5).$$

The pivots are in columns 2 and 4, so x_2, x_4 are pivotal and x_1, x_3, x_5 are free. (Column 1 is entirely zero, so x_1 does not appear in any equation at all; it is free.) The two nonzero rows say $x_2 + 2x_3 + 3x_5 = 4$ and $x_4 - 2x_5 = -1$. Solving for the pivotal variables:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} x_1 \\ 4 - 2x_3 - 3x_5 \\ x_3 \\ -1 + 2x_5 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 0 \\ -1 \\ 0 \end{bmatrix} + x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ -3 \\ 0 \\ 2 \\ 1 \end{bmatrix}.$$

Each vector is read off by collecting terms: the constants give $(0, 4, 0, -1, 0)$; the coefficients of x_1, x_3, x_5 give the other three. The d column has no pivot and columns 1, 3, 5 are not pivotal, so by the theorem there are infinitely many solutions (a 3-parameter family).

Check: the nonzero rows of R send $(0, 4, 0, -1, 0)$ to $(4, -1) = d$, and send each of the three direction vectors to $(0, 0)$. ✓

Page 12 ("Is it in RREF?" examples) is filed with the 8/25 notes, not here; see the 8/25 notes for it. Its three examples: $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not in RREF (the pivot in column 2 has a nonzero entry above it), $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is not (the zero row is not at the bottom), and $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is.

3 Interpolating polynomials (p. 13)

Finding a polynomial through given points is a linear system whose unknowns are the coefficients.

Goal. Draw the graph of a degree- n polynomial

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

through given data points $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$. Requiring $p(x_i) = y_i$ for each point gives

$$\begin{cases} a_n x_0^n + \dots + a_1 x_0 + a_0 = y_0 \\ \vdots \\ a_n x_n^n + \dots + a_1 x_n + a_0 = y_n. \end{cases}$$

The x_i, y_i are known numbers, so each equation is *linear* in the unknowns $a_n, \dots, a_0$ (the powers x_i^k are just coefficients). This is a linear system with $n + 1$ equations and $n + 1$ unknowns $a_n, \dots, a_0$. (The notes index the points $1, \dots, n$; see the corrections table.)

Example (p. 13): the parabola through $(1, 2), (2, 0), (-1, 1)$

Take $p(x) = ax^2 + bx + c$. The three conditions $p(1) = 2, p(2) = 0, p(-1) = 1$ are

$$\begin{cases} a \cdot 1^2 + b \cdot 1 + c = 2 \\ a \cdot 2^2 + b \cdot 2 + c = 0 \\ a(-1)^2 + b(-1) + c = 1. \end{cases}$$

“Can find a, b, c by row reduction”: $p(x) = -\frac{5}{6}x^2 + \frac{1}{2}x + \frac{7}{3}$

The augmented matrix (columns a, b, c and right-hand side) and its row reduction:

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 4 & 2 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{array} \right] &\xrightarrow[R_3 - R_1]{R_2 - 4R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -2 & -3 & -8 \\ 0 & -2 & 0 & -1 \end{array} \right] \xrightarrow{R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -2 & -3 & -8 \\ 0 & 0 & 3 & 7 \end{array} \right] \\ &\xrightarrow[\frac{1}{3}R_3]{-\frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{3}{2} & 4 \\ 0 & 0 & 1 & \frac{7}{3} \end{array} \right] \xrightarrow{\dots} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{5}{6} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{7}{3} \end{array} \right]. \end{aligned}$$

The last step clears the entries above the pivots: $R_2 - \frac{3}{2}R_3$ gives $b = 4 - \frac{7}{2} = \frac{1}{2}$, then $R_1 - R_2 - R_3$ gives $a = 2 - \frac{1}{2} - \frac{7}{3} = -\frac{5}{6}$. Every column of the coefficient part is pivotal and the last column is not, so by the theorem of p. 10 there is exactly one such parabola:

$$a = -\frac{5}{6}, \quad b = \frac{1}{2}, \quad c = \frac{7}{3}.$$

Check: $p(1) = \frac{-5+3+14}{6} = 2, p(2) = \frac{-20+6+14}{6} = 0, p(-1) = \frac{-5-3+14}{6} = 1. \checkmark$

4 Vectors in $\mathbb{R}^n$ (pp. 14–15)

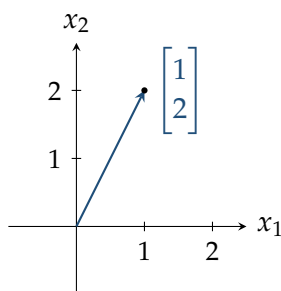
A vector in $\mathbb{R}^n$ is a column of n real numbers; vectors can be added and scaled, and addition follows the parallelogram rule.

Definition (p. 14)

$\mathbb{R}^n$ is the set of all matrices $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ (columns with n entries) where each x_i is a real number.

Notation: $\mathbb{R}^1 = \mathbb{R}$.

Lay's text makes a distinction between a *point* in $\mathbb{R}^n$, written $(x_1, \dots, x_n)$, and a *vector*, an arrow from the origin $(0, \dots, 0)$ to the point $(x_1, \dots, x_n)$, written as a column. For example ($n = 2$), the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is the arrow from the origin to the point $(1, 2)$:



We ignore the distinction and write points and vectors in whichever form is convenient. In running text these notes often write a column vector as $(x_1, \dots, x_n)$ or $(x_1, \dots, x_n)^T$ to save space.

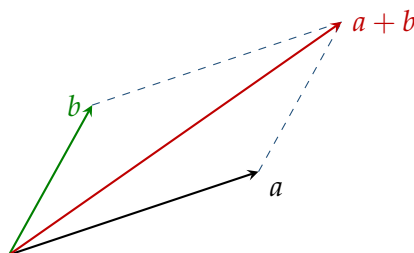
Informal definition. Vectors are things that can be added and scaled. We shall see more examples later when we talk about vector spaces. In $\mathbb{R}^n$ both operations are entry by entry:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix}, \quad c \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} cu_1 \\ \vdots \\ cu_n \end{bmatrix} \quad (\text{Lay §1.3}).$$

(These formulas are not written on the page but are used on every page that follows.)

The parallelogram rule (p. 15)

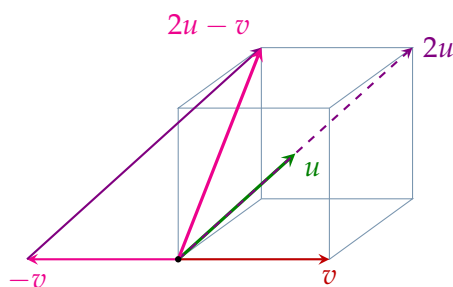
To add a and b geometrically, place the tail of b at the tip of a (or the other way round). The sum $a + b$ is the diagonal of the parallelogram with sides a and b :



The figure shows a , b , and the sum $a + b$ as the long diagonal. The dashed sides are copies of a and b moved parallel to themselves.

Example (p. 15): cube diagonal

The sketch on p. 15 shows a box with an edge vector v (red), a vector u (green) pointing into the box, and the combination $2u - v$ (pink), built with the parallelogram rule as $2u + (-v)$: first go backwards along $-v$, then add $2u$. The sketch is rough; *my reading* is that u is half the main diagonal of the cube, so that $2u$ is the cube diagonal (hence the title). Redrawn with a cube of side 2, $v = (2, 0, 0)$ and $u = (1, 1, 1)$:



The figure shows the cube, v and $-v$ along the bottom front edge, u and the diagonal $2u$, and $2u - v$ obtained by attaching $2u$ (solid violet) to the tip of $-v$. Here $2u - v = (2, 2, 2) - (2, 0, 0) = (0, 2, 2)$, the diagonal of the left face. Check: $-v + 2u = (-2, 0, 0) + (2, 2, 2) = (0, 2, 2)$. ✓

5 Linear combinations and span (pp. 15–17)

The span of some vectors is everything you can reach from them by scaling and adding; deciding whether b is in the span is solving a linear system.

Key technical notions.

Definition (p. 15; Lay §1.3)

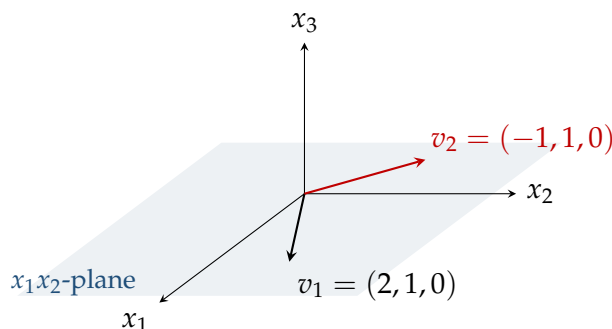
Given vectors $v_1, \dots, v_n$ and numbers $c_1, \dots, c_n$, the vector

$$v = c_1 v_1 + \dots + c_n v_n$$

is a **linear combination** of $v_1, \dots, v_n$ with **weights** $c_1, \dots, c_n$.

The set of *all* linear combinations of $v_1, \dots, v_n$ is $\text{span}\{v_1, \dots, v_n\}$, the **span** of $v_1, \dots, v_n$.

Example (pp. 16–17): $v_1 = (2, 1, 0)$, $v_2 = (-1, 1, 0)$ in $\mathbb{R}^3$



The figure shows both vectors lying in the shaded x_1x_2 -plane (both have third entry 0). The notes ask: **what is $\text{span}\{v_1, v_2\}$?** A vector $b = (b_1, b_2, b_3)$ is in the span when

$$c_1 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b$$

for some weights c_1, c_2 . We must describe all such b .

What can we say about b_3 ? It must be 0

The third entry of the left side is $c_1 \cdot 0 + c_2 \cdot 0 = 0$ for every choice of weights. So every b in the span has $b_3 = 0$; the span lies inside the x_1x_2 -plane.

So we actually have the system (dropping the third entries)

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad \text{i.e.} \quad \begin{cases} 2c_1 - c_2 = b_1 \\ c_1 + c_2 = b_2. \end{cases}$$

Is every (b_1, b_2) realized for some c_1, c_2 ? Yes

Row reduce the coefficient matrix (the notes swap the rows, then subtract twice row 1 from row 2):

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 1 \\ 0 & -3 \end{bmatrix}.$$

Two pivots, so yes. The reason, given in the Note on p. 17: whatever b is, row reducing $[A \mid b]$ gives $[R \mid d]$ where R has a pivot in each of its two rows, so there is no room for a pivot in the d column, and by the theorem of p. 10 the system has a solution.

Explicitly, adding the two equations gives $3c_1 = b_1 + b_2$, so

$$c_1 = \frac{b_1 + b_2}{3}, \quad c_2 = b_2 - c_1 = \frac{2b_2 - b_1}{3}.$$

Check: $2c_1 - c_2 = \frac{2b_1 + 2b_2 - 2b_2 + b_1}{3} = b_1$ and $c_1 + c_2 = \frac{3b_2}{3} = b_2$. ✓

Note (p. 17)

Since row reduction proceeds from the left, the RREF of $[A \mid b]$ is $[\text{RREF}(A) \mid d]$ for some column d . In particular, if $\text{RREF}(A) = \begin{bmatrix} 1 & 0 \\ \cdot & \cdot \\ 0 & 1 \end{bmatrix}$ (the identity), then there is no pivot in the rightmost column of $\text{RREF}[A \mid b]$, so $Ax = b$ has a solution.

Fill-in: $\text{RREF}[A \mid b] = [\text{RREF}(A) \mid d]$, where d is whatever b becomes

The “?” is a column d that depends on b . The row reduction algorithm (8/25 notes, pp. 7–8) works on the leftmost nonzero column first, and its choices (which row to swap up, what to scale and subtract) depend only on the entries in the columns it has reached. So the first n columns of $[A \mid b]$ are reduced exactly as A alone would be, giving $\text{RREF}(A)$; the last column just gets carried along, ending as some d .

Why no pivot on the right when $\text{RREF}(A) = I$. Every row of I already has its pivot on the left, and a pivot must be the first nonzero entry of its row, so the last column cannot contain one. By the theorem of p. 10, part (1), $Ax = b$ is consistent; in fact $x = d$.

For v_1, v_2 above: $\text{RREF} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = I$, since one more step $-\frac{1}{3}R_2$, then $R_1 - R_2$, turns $\begin{bmatrix} 1 & 1 \\ 0 & -3 \end{bmatrix}$ into I .

What is $\text{span}\{v_1, v_2\}$? The x_1x_2 -plane

The first box shows the span lies in the x_1x_2 -plane $\{(b_1, b_2, 0)\}$. The second shows every $(b_1, b_2, 0)$ is reached. So $\text{span}\{v_1, v_2\} = \{(b_1, b_2, 0) : b_1, b_2 \in \mathbb{R}\}$, the x_1x_2 -plane, which is what the notes conclude on p. 17.

Check with $b = (1, 0, 0)$: $c_1 = \frac{1}{3}$, $c_2 = -\frac{1}{3}$, and $\frac{1}{3}(2, 1, 0) - \frac{1}{3}(-1, 1, 0) = (1, 0, 0)$. ✓

Two more spans (p. 17)

- If v is a nonzero vector in $\mathbb{R}^n$, then $\text{span}\{v\} = \{cv : c \in \mathbb{R}\}$ is the **line through the origin in the direction of v** .
- If $v = \mathbf{0} = (0, \dots, 0)$, then $\text{span}\{v\} = \{\mathbf{0}\}$, since $c \cdot \mathbf{0} = \mathbf{0}$ for every c . The span is a single point, not a line; this is why the first example needs $v \neq \mathbf{0}$.

6 The matrix–vector product (p. 18)

Ax is defined as the linear combination of the columns of A with weights the entries of x .

Goal: write $Ax = b$ as a vector equation.

Definition (matrix–vector product, p. 18; Lay §1.4)

Let A be an $m \times n$ matrix (m rows, n columns) with columns $a_1, \dots, a_n$ in $\mathbb{R}^m$, and let $x =$

$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ be a vector in $\mathbb{R}^n$. Then

$$Ax = [a_1 \cdots a_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \stackrel{\text{def}}{=} x_1 a_1 + \cdots + x_n a_n,$$

the linear combination of $a_1, \dots, a_n$ with weights $x_1, \dots, x_n$.

Writing $A = [a_1 \cdots a_n]$ just means listing A by its columns. The number of entries of x must equal the number of columns of A , and the result has as many entries as A has rows.

Note. $\text{span}\{a_1, \dots, a_n\} = \{Ax : x \in \mathbb{R}^n\}$. Indeed, as x runs over $\mathbb{R}^n$, the weights $x_1, \dots, x_n$ run over all possible choices, so Ax runs over all linear combinations of the columns.

Example (p. 18)

Sizes: a 2×3 matrix times a 3×1 vector gives a 2×1 vector.

Fill-in: $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \end{bmatrix} (4, 3, 7)^T = (-3, -8)^T$

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + 7 \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 + 0 - 7 \\ 0 + 6 - 14 \end{bmatrix} = \begin{bmatrix} -3 \\ -8 \end{bmatrix}.$$

Check with the row-times-column rule (Lay §1.4): row 1 gives $1 \cdot 4 + 0 \cdot 3 + (-1) \cdot 7 = -3$, row 2 gives $0 \cdot 4 + 2 \cdot 3 + (-2) \cdot 7 = -8$. ✓

Q (p. 18): the span of the columns of $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

The columns $(1, 4)$, $(2, 5)$, $(3, 6)$ have two entries, so their span lies in $\mathbb{R}^2$. The natural question is whether it is *all* of $\mathbb{R}^2$.

Is the span of the columns all of $\mathbb{R}^2$? Yes

Row reduce: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \xrightarrow{R_2 - 4R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \end{bmatrix} \xrightarrow{-\frac{1}{3}R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{R_1 - 2R_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$. There is a pivot in every row, so by the theorem of p. 19 (Section 6 below) $Ax = b$ is consistent for every $b \in \mathbb{R}^2$, i.e. the columns span $\mathbb{R}^2$. In fact the first two columns alone already do (same argument as p. 16).

Check: $(1, 0)$ is reached: $-\frac{5}{3}(1, 4) + \frac{4}{3}(2, 5) = (\frac{-5+8}{3}, \frac{-20+20}{3}) = (1, 0)$, and $(0, 1)$ is reached: $\frac{2}{3}(1, 4) - \frac{1}{3}(2, 5) = (0, 1)$. Every $(b_1, b_2) = b_1(1, 0) + b_2(0, 1)$ is then a combination too. ✓

7 $Ax = b$ as a vector equation; when is it always solvable? (pp. 19–21)

$Ax = b$ has a solution exactly when b is in the span of the columns of A , and it has a solution for every b exactly when A has a pivot in every row.

Thus $Ax = b$ can be written as the **vector equation**

$$x_1a_1 + \cdots + x_na_n = b.$$

Theorem (p. 19)

$Ax = b$ is consistent (has a solution) $\iff b$ is in the span of the columns $a_1, \dots, a_n$ of A .

Proof (“trivial”), written out

$Ax = b$ has a solution $x = (x_1, \dots, x_n)$ $\iff$ there are numbers $x_1, \dots, x_n$ with $x_1a_1 + \cdots + x_na_n = b$ (definition of Ax , p. 18) $\iff b$ is a linear combination of $a_1, \dots, a_n$ $\iff b \in \text{span}\{a_1, \dots, a_n\}$ (definition of span, p. 15). $\square$

Theorem (p. 19; Lay §1.4, Theorem 4)

Let A be an $m \times n$ matrix. The following are equivalent:

1. $Ax = b$ is consistent for every b in $\mathbb{R}^m$.
2. The columns of A span $\mathbb{R}^m$.
3. A has a pivot in every row (that is, $\text{RREF}(A)$ has m pivots).

“ A has a pivot in every row” means $\text{RREF}(A)$ does; since each row holds at most one pivot, that is the same as m pivots.

Proof of 1 $\iff$ 2 (“from definitions”)

By the previous theorem, $Ax = b$ is consistent exactly when $b \in \text{span}\{a_1, \dots, a_n\}$. So “consistent for every $b \in \mathbb{R}^m$ ” says “every $b \in \mathbb{R}^m$ lies in $\text{span}\{a_1, \dots, a_n\}$ ”, which is what “the columns span $\mathbb{R}^m$ ” means (the span is always inside $\mathbb{R}^m$, so the two sets are equal). $\square$

Proof of 3 $\Rightarrow$ 1 and 1 $\Rightarrow$ 3 (pp. 20–21), with the details filled in

3 $\Rightarrow$ 1. Fix any b . We wish to show that $\text{RREF}[A \mid b]$ has no pivot in the rightmost column; then $Ax = b$ is consistent by the theorem of p. 10, part (1). Row reduction proceeds from the left, so $\text{RREF}[A \mid b] = [\text{RREF}(A) \mid d]$ for some d (Note, p. 17). If $\text{RREF}(A)$ has a pivot in each row, every row’s first nonzero entry is already to the left of the bar. So there cannot be a pivot in the d column.

1 $\Rightarrow$ 3. Argue by contradiction. Suppose $Ax = b$ is consistent for each b , but $R := \text{RREF}(A)$ does not have a pivot in every row. Then the bottom row of R has no pivot (pivots move down and to the right, and zero rows are at the bottom, 8/25 notes, p. 7). A row of an RREF

matrix with no pivot has no leading entry, so it is entirely zero. Now let

$$d = e_m = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \text{so} \quad [R \mid d] = \left[\begin{array}{ccc|c} * & \cdots & * & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{array} \right].$$

The bottom row says $0 = 1$, so $Rx = d$ is inconsistent.

Take the row operations that took A to R , and do them in reverse: undo the last one first, using the inverse operations (each row operation is reversible, 8/25 notes, p. 6). Apply these to $[R \mid d]$. On the left part they turn R back into A ; the right column becomes some vector b that depends on d . So we arrive at $[A \mid b]$.

Since $[A \mid b]$ and $[R \mid d]$ are related by row operations, $Ax = b$ and $Rx = d$ have the same solutions (8/25 notes, p. 6). Since $Rx = d$ is inconsistent, $Ax = b$ is inconsistent. This contradicts 1. $\square$

Example (p. 21)

$$A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \end{bmatrix}. \text{ Is } Ax = b \text{ consistent for every } b?$$

Is $Ax = b$ consistent for every $b \in \mathbb{R}^2$? Yes

Row reduce:

$$\left[\begin{array}{ccc} 1 & 2 & 0 \\ -1 & 1 & 2 \end{array} \right] \xrightarrow{R_2+R_1} \left[\begin{array}{ccc} 1 & 2 & 0 \\ 0 & 3 & 2 \end{array} \right] \xrightarrow{\frac{1}{3}R_2} \left[\begin{array}{ccc} 1 & 2 & 0 \\ 0 & 1 & \frac{2}{3} \end{array} \right] \xrightarrow{R_1-2R_2} \left[\begin{array}{ccc} 1 & 0 & -\frac{4}{3} \\ 0 & 1 & \frac{2}{3} \end{array} \right].$$

A has a pivot in both rows, so by the theorem ($3 \Rightarrow 1$) $Ax = b$ is consistent for every b , and the columns of A span $\mathbb{R}^2$. (x_3 is free, so by the theorem of p. 10 there are in fact infinitely many solutions for each b .)

One explicit solution: take the free variable $x_3 = 0$; then $x_1 + 2x_2 = b_1$ and $-x_1 + x_2 = b_2$ give

$$x = \left(\frac{b_1 - 2b_2}{3}, \frac{b_1 + b_2}{3}, 0 \right).$$

Check: $x_1 + 2x_2 = \frac{b_1 - 2b_2 + 2b_1 + 2b_2}{3} = b_1$, and $-x_1 + x_2 + 2 \cdot 0 = \frac{-b_1 + 2b_2 + b_1 + b_2}{3} = b_2$. $\checkmark$

Pitfall. The test is a pivot in every *row* of A , not every column. A pivot in every column is the uniqueness condition (2a) of p. 10. For example $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ has a pivot in its only column, yet $x \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ has no solution. Also, the test must be applied to A , not to $[A \mid b]$: a pivot in the last column of $[A \mid b]$ is exactly what makes the system inconsistent.

Summary

Fact	Formula / test
No solutions (p. 10)	pivot in the d column of $[R \mid d]$
Unique solution (p. 10)	consistent, and every column of R pivotal
Infinitely many (p. 10)	consistent, and some column of R not pivotal (a free variable)
Parametric form (p. 9)	$x = p + t_1 w_1 + \cdots + t_k w_k$, $t_i = \text{free variables}$
Matrix–vector product (p. 18)	$Ax = x_1 a_1 + \cdots + x_n a_n$
Span (p. 18)	$\text{span}\{a_1, \dots, a_n\} = \{Ax : x \in \mathbb{R}^n\}$
Consistency (p. 19)	$Ax = b$ solvable $\iff b \in \text{span}\{a_1, \dots, a_n\}$
Solvable for all b (p. 19)	columns span $\mathbb{R}^m \iff$ pivot in every row of A
Interpolation (p. 13)	$p(x_i) = y_i$ is linear in the coefficients $a_n, \dots, a_0$

Corrections to the handwritten notes

Page	As written	Should be	Why
10	“a linear system on RREF of $[A \mid b]$ is $[R \mid d]$ ”	“...system <i>and</i> RREF ...”	word slip
13	data points $(x_1, y_1), \dots, (x_n, y_n)$; “ n equations and unknowns $a_n \dots a_0$ ”	$n + 1$ points $(x_0, y_0), \dots, (x_n, y_n)$; $n + 1$ equations	$a_n, \dots, a_0$ are $n + 1$ unknowns; the example (degree 2) uses 3 points
20	“So $Ax = b$ is consistent for each b ”	“Suppose $Ax = b$ is consistent ...”	start of a proof by contradiction (an assumption, not a conclusion)