

Linear Algebra Notes — Lecture 9/15

The Invertible Matrix Theorem, LU Factorization, Subspaces of $\mathbb{R}^n$

Lecture pages 75–89 · Lay §2.3, §2.5, §2.8

Contents

Overview	1
1 Plan, and how MATLAB computes A^{-1} (p. 75)	2
2 The Invertible Matrix Theorem (slide 48; p. 75)	2
3 Triangular, diagonal and permutation matrices (pp. 76–78)	4
4 The permuted LU factorization (pp. 78–80)	6
5 How to compute $A = LU$ (pp. 81–83)	7
6 Subspaces of $\mathbb{R}^n$ (pp. 84–87)	9
7 Bases, and a basis of $\text{Nul } A$ (pp. 87–89)	11
Summary	13
Corrections to the handwritten notes	13

Overview

Last time (9/10) we built inverses: the algorithm $[A \ I_n] \rightarrow [I_n \ A^{-1}]$, elementary matrices, and left and right inverses. This lecture does three things:

1. **The Invertible Matrix Theorem (IMT).** Twelve statements about a square matrix are all equivalent. Almost every link was already proved in earlier lectures; we collect the proofs.
2. **LU factorization.** Row reduction can be recorded as $A = PLU$ with L lower triangular and U an echelon form. Then each new system $Ax = b$ costs only two cheap triangular solves.
3. **Subspaces of $\mathbb{R}^n$.** Sets closed under addition and scaling: null spaces, column spans, lines through 0. A *basis* is an independent spanning set, and the usual way of solving $Ax = 0$ gives a basis of $\text{Nul } A$.

The test is next Wednesday (written in red on p. 75).

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

Notation used throughout.

- e_i : column i of I_n (a 1 in slot i , zeros elsewhere). Ae_i is column i of A (9/10 notes, p. 74).
- $T(x) = Ax$: the linear map with standard matrix A .
- $O(n^k)$: “at most a constant times n^k ” operations. A *flop* is one floating-point operation ($+$, $-$, $\times$, $\div$).
- $\text{Nul } A = \{x : Ax = 0\}$, $\text{Ker } T$, $\text{Col } A$: defined in Section 6.
- $T^{-1}(0) = \{x : T(x) = 0\}$: the *preimage* of 0. This notation makes sense even when T has no inverse.

1 Plan, and how MATLAB computes A^{-1} (p. 75)

Inverting with $[A \ I] \rightarrow [I \ A^{-1}]$ costs $O(n^3)$ flops. MATLAB does the same work by factoring $A = PLU$ first.

The notes write

$$[A \ I] \longrightarrow [I \ A^{-1}] \quad O(n^3), \quad A = PLU, \quad A^{-1} = U^{-1}L^{-1}P^{-1}.$$

The last formula is the rule $(XYZ)^{-1} = Z^{-1}Y^{-1}X^{-1}$ (9/10 notes, p. 64–65), which reverses the order. Each of P , L , U is easy to invert (my reading of the three small arrows under $U^{-1}L^{-1}P^{-1}$ is that they point at these three easy pieces). The matrices P , L , U are defined in Sections 3–4.

In practice one rarely forms A^{-1} at all. To solve $Ax = b$, MATLAB uses $A = PLU$ directly (Section 4). This is the “Why?” on p. 67 of the 9/10 notes: solving $Ax = b$ directly is cheaper than computing A^{-1} first.

2 The Invertible Matrix Theorem (slide 48; p. 75)

For a square matrix, “invertible”, “ n pivots”, “one-to-one”, “onto” and nine other statements are all the same condition.

Theorem (Invertible Matrix Theorem, Lay §2.3, Theorem 8)

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

- A is invertible.
- A is row equivalent to I_n .
- A has n pivotal columns. (All columns are pivotal.)
- $Ax = 0$ has only the trivial solution.

- e) The columns of A are linearly independent.
- f) The linear transformation $x \mapsto Ax$ is one-to-one.
- g) The equation $Ax = b$ has a solution for every $b \in \mathbb{R}^n$.
- h) The columns of A span $\mathbb{R}^n$.
- i) The linear transformation $x \mapsto Ax$ is onto.
- j) There is an $n \times n$ matrix C with $CA = I_n$. (A has a left inverse.)
- k) There is an $n \times n$ matrix D with $AD = I_n$. (A has a right inverse.)
- l) A^T is invertible.

The notes say “see separate file on IMT and why we have already proven it”. The diagram on the slide is the plan of the proof:

$$a \Leftrightarrow b \text{ (last time),} \quad b \Leftrightarrow c \Leftrightarrow d \Leftrightarrow e, \quad d \Leftrightarrow f, \quad b \Leftrightarrow g \Leftrightarrow h \Leftrightarrow i, \\ g \Leftrightarrow \text{“all rows are pivotal”}, \quad j \Rightarrow b, k \Rightarrow b \text{ (homework 3)}, \quad a \Rightarrow j, a \Rightarrow k, \quad a \Leftrightarrow l.$$

Since every statement connects to (a) or (b), all twelve are equivalent.

Proof of the IMT: every link in the diagram, filled in

- **(a) $\Leftrightarrow$ (b).** Theorem on p. 66 of the 9/10 notes (A invertible $\Leftrightarrow \text{RREF}(A) = I_n$), proved on pp. 67–73.
- **(b) $\Leftrightarrow$ (c).** The RREF of a square matrix with a pivot in every column has a 1 on each diagonal spot and 0 elsewhere, so it is I_n . Conversely I_n has n pivots.
- **(c) $\Leftrightarrow$ (d).** $Ax = 0$ has a nontrivial solution exactly when there is a free variable, i.e. a non-pivotal column (Lay §1.5).
- **(d) $\Leftrightarrow$ (e).** Write $A = [a_1 \cdots a_n]$. The notes’ line is

$$c_1 a_1 + \cdots + c_n a_n = A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

So a dependence relation $c_1 a_1 + \cdots + c_n a_n = 0$ with some $c_i \neq 0$ is the same thing as a nontrivial solution of $Ax = 0$.

- **(d) $\Leftrightarrow$ (f).** With $T(x) = Ax$ (the notes’ line): if $T(u) = T(v)$, then by linearity $T(u - v) = A(u - v) = 0$, so (d) gives $u - v = 0$, i.e. $u = v$. Conversely, if T is one-to-one and $Ax = 0 = A0$, then $x = 0$.
- **(b) $\Leftrightarrow$ (g).** (g) holds exactly when A has a pivot in every row (9/8 notes, p. 49; Lay §1.4, Theorem 4). This is the red remark “g $\Leftrightarrow$ rows are pivotal”. A square matrix has a pivot in every row iff it has n pivots iff it has a pivot in every column, which is (c), hence (b).
- **(g) $\Leftrightarrow$ (h) $\Leftrightarrow$ (i).** $Ax = x_1 a_1 + \cdots + x_n a_n$, so “ $Ax = b$ is solvable” means “ b is a combination of the columns”. Hence (g) says every b is in $\text{span}\{a_1, \dots, a_n\}$, which is (h). And “ T onto” means every b equals $T(x) = Ax$ for some x , which is (g) word for word.

- **(j)⇒(b)** (homework 3). If $CA = I_n$ and $Ax = 0$, then $x = I_n x = CAx = C0 = 0$. So (d) holds, hence (b).
- **(k)⇒(b)** (homework 3). If $AD = I_n$, then for any b the vector $x = Db$ solves $Ax = ADb = b$. So (g) holds, hence (b).
- **(a)⇒(j), (a)⇒(k)**. Take $C = D = A^{-1}$.
- **(a)⇔(l)**. If A is invertible, so is A^T , with $(A^T)^{-1} = (A^{-1})^T$ (the notes' last line; 9/10 notes, pp. 64–65): indeed $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$. Conversely, if A^T is invertible, apply the same fact to A^T : $(A^T)^T = A$ is invertible.

Note (p. 75): A invertible $\Leftrightarrow T$ one-to-one and onto $\Leftrightarrow T$ invertible as a function. Why?

The first $\Leftrightarrow$ is (a)⇔(f) and (a)⇔(i). For the second: a function is invertible exactly when it is one-to-one and onto (each b comes from exactly one x). If A is invertible, the inverse function is $T^{-1}(b) = A^{-1}b$, since $A(A^{-1}b) = b$ and $A^{-1}(Ax) = x$.

Example: $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ has 2 pivots, so T is one-to-one and onto; $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ has 1 pivot, so T is neither (e.g. $T(2, -1) = T(0, 0) = 0$). ✓

Square matrices only. For a non-square matrix “one-to-one” and “onto” are different conditions. The left-inverse example on p. 74 of the 9/10 notes, $A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, has $CA = I_1$ with $C = [1 \ x]$, but $AC \neq I_2$. For an $n \times n$ matrix the IMT says one side is enough: $CA = I$ forces $AC = I$ (and then $C = A^{-1}$).

3 Triangular, diagonal and permutation matrices (pp. 76–78)

Products of lower triangular matrices are lower triangular, and permutation matrices just reorder the entries of a vector.

Definition (p. 76)

An $n \times n$ matrix $A = (a_{ij})$ is

- **upper triangular** if $a_{ij} = 0$ for all $i > j$ (zeros below the diagonal);
- **lower triangular** if $a_{ij} = 0$ for all $i < j$ (zeros above the diagonal).

A matrix that is both upper and lower triangular is **diagonal**: $\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}$.

Examples (p. 76).

$$\text{Upper: } \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{bmatrix}; \quad \text{Lower: } \begin{bmatrix} 0 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 \\ 2 & 3 & 0 \\ 1 & -1 & 0 \end{bmatrix}.$$

In the notes a red line marks the diagonal of the second example: everything below it is 0. Diagonal entries may be zero; only the entries on the wrong side of the diagonal must vanish.

Example (pp. 76–77): product of lower triangular matrices

Let A be lower triangular with diagonal $d_1, \dots, d_n$ and B lower triangular with diagonal $c_1, \dots, c_n$ (entries marked $*$ are arbitrary).

Fill-in: if $i < j$, the (i, j) entry of AB is 0

The (i, j) entry of AB is (row i of A) · (column j of B) (9/8 notes, p. 54):

$$\begin{bmatrix} * & \cdots & * & d_i & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ c_j \\ * \\ \vdots \\ * \end{bmatrix}.$$

Row i of A is zero after position i . Column j of B is zero before position j . If $i < j$, every position k has $k > i$ (so $a_{ik} = 0$) or $k < j$ (so $b_{kj} = 0$), because there is no k with $k \leq i$ and $k \geq j$. So every term $a_{ik}b_{kj}$ is 0, and the entry is 0. Hence AB is **lower triangular**. Bonus: the diagonal entries of AB are d_ic_i (only $k = i$ survives).

Fill-in (p. 77): $\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 1 & 3 \end{bmatrix}$

The notes only show the shape $\begin{bmatrix} * & 0 \\ * & * \end{bmatrix}$; the red marks show the $(1, 2)$ entry as row 1 of the left factor times column 2 of the right factor: $1 \cdot 0 + 0 \cdot 1 = 0$. The other entries: $(1, 1) = 1(-1) + 0(1) = -1$, $(2, 1) = 2(-1) + 3(1) = 1$, $(2, 2) = 2 \cdot 0 + 3 \cdot 1 = 3$. *Check:* the diagonal is $1 \cdot (-1) = -1$ and $3 \cdot 1 = 3$, the products of the diagonals, as the general argument predicts. ✓

Similarly, a product of upper triangular matrices is upper triangular. Why?

Same argument with the roles reversed: if $i > j$, row i of A is zero before position i and column j of B is zero after position j , so no term survives. Or use transposes: A, B upper $\Rightarrow A^T, B^T$ lower $\Rightarrow B^T A^T = (AB)^T$ lower (9/10 notes, p. 61) $\Rightarrow AB$ upper.

Definition (permutation matrix, p. 77)

An $n \times n$ matrix P is a **permutation matrix** if

$$P \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{i_1} \\ \vdots \\ x_{i_n} \end{bmatrix} \quad \text{for all } x, \quad \text{where } \{i_1, \dots, i_n\} = \{1, \dots, n\}.$$

So P reorders the entries of x ; each entry is used exactly once.

Fill-in: the 2×2 permutation matrices are exactly $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

There are only two orderings of (x_1, x_2) : (x_2, x_1) (swap) and (x_1, x_2) (do nothing). Since the columns of P are Pe_1, Pe_2 , the swap gives $[e_2 \ e_1]$ and the identity gives I_2 . In general there are $n!$ permutation matrices of size n , one for each ordering.

Example (p. 78): a 3×3 permutation

We want P with $P(x_1, x_2, x_3) = (x_2, x_3, x_1)$. As for any linear map, column i of P is $P(e_i)$ (9/8 notes, p. 49). Putting $x = e_1 = (1, 0, 0)$ gives $(0, 0, 1)$; $x = e_2$ gives $(1, 0, 0)$; $x = e_3$ gives $(0, 1, 0)$. So

$$P = [P(e_1) \ P(e_2) \ P(e_3)] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Check: $P(x_1, x_2, x_3)^T = (0x_1 + 1x_2 + 0x_3, \ x_3, \ x_1)^T = (x_2, x_3, x_1)^T$. ✓

Meaning: every row and every column of a permutation matrix has exactly one 1. Row k has its 1 in column i_k , the position whose entry moves to slot k .

4 The permuted LU factorization (pp. 78–80)

Every matrix is a permutation, times a lower triangular matrix with 1s on the diagonal, times an echelon form; this turns solving $Ax = b$ into two triangular solves.

Theorem (permuted LU factorization, p. 78; Lay §2.5)

If A is an $m \times n$ matrix, then $A = PLU$, where

- P is an $m \times m$ permutation matrix;
- $L = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ * & & 1 \end{bmatrix}$ is $m \times m$ lower triangular with 1s on the diagonal;
- U is an echelon form (REF) of A , e.g. $\begin{bmatrix} \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{bmatrix}$ with pivots $\blacksquare \neq 0$ and $*$ any number.

U is the intermediate step on the way from A to RREF: zeros below the pivots, but pivots not yet scaled to 1 and no zeros above them. U is “upper triangular” in the echelon sense (a square U is upper triangular). The permutation P is needed only when row swaps are; see Section 5.

Motivation (p. 79)

Suppose we solve $Ax = b_1, Ax = b_2, \dots, Ax = b_k$: one A , many right-hand sides. Reducing each $[A \ b_i]$ to RREF repeats the same row operations on A every time. **Goal:** reuse these row operations. The matrix L is exactly the stored record of them.

Use of the LU factorization (pp. 79–80)

Given $A = PLU$, solve $Ax = b$ by splitting it into two systems:

$$(1) Ux = y, \quad (2) Ly = P^{-1}b,$$

where P^{-1} is the inverse permutation matrix (it undoes the reordering).

Why do (1) and (2) solve $Ax = b$? And in which order?

The notes' line: $b = Ax = PL \underbrace{Ux}_y$. Name $y = Ux$. Then $b = PLy$, and multiplying by P^{-1} gives $Ly = P^{-1}b$. So: **first** solve (2) for y (it only involves known data L, P, b), **then** solve (1) for x . Conversely, if $Ly = P^{-1}b$ and $Ux = y$, then $Ax = PLUx = PLy = PP^{-1}b = b$. ✓
 P^{-1} is again a permutation matrix and in fact $P^{-1} = P^T$. For the p. 78 example, $P^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ sends (x_2, x_3, x_1) back to (x_1, x_2, x_3) , and one checks $P^T P = I_3$. ✓

Key point (p. 80). U and L are triangular, so $Ux = y$ and $Ly = P^{-1}b$ are easy to solve: by substitution, with no row reduction.

Facts (p. 80). If A is $n \times n$, then

- finding $A = PLU$ takes $O(n^3)$ flops (flop = operation of floating-point arithmetic; $O(n^3)$ = a constant times n^3);
- solving $Ax = b$ (by row reduction) takes $O(n^3)$ flops;
- solving $Ux = y$ and $Ly = P^{-1}b$ takes $O(n^2)$ flops.

Why $O(n^2)$: in back substitution, finding x_k uses the $\leq n$ already-known later unknowns, so about n operations per unknown, $n \cdot n$ in total. Reducing a full matrix, by contrast, updates up to n^2 entries for each of n pivots. So for k right-hand sides, LU costs $O(n^3) + k \cdot O(n^2)$ instead of $k \cdot O(n^3)$.

Example (pp. 80–81): a triangular system solved by substitution

Fill-in: $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ has $x_1 = x_2 = \frac{1}{3}$

The last row says $3x_2 = 1$, so $x_2 = \frac{1}{3}$. Substituting into the first row $2x_1 + x_2 = 1$ gives $x_1 = \frac{1 - \frac{1}{3}}{2} = \frac{1}{3}$. No row reduction needed. Check: $2 \cdot \frac{1}{3} + \frac{1}{3} = 1$ and $3 \cdot \frac{1}{3} = 1$. ✓

5 How to compute $A = LU$ (pp. 81–83)

If A reduces to echelon form using only row replacements, then L is the record of the multipliers and $P = I$.

Practical answer (p. 81): use MATLAB. The idea behind it:

Simplifying assumption. A can be reduced to RREF (equivalently, to REF) without row swaps. Then $P = I$ and $A = LU$.

Algorithm (p. 81)

1. Reduce A to an echelon form U using row replacement only (no scaling).
2. Place the row-replacement factors into L : if the step is “row $i \leftarrow \text{row } i + c \cdot \text{row } j$ ”, put $-c$ in position (i, j) of L . Put 1s on the diagonal and 0s above it.

Example (pp. 81–82)

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & -1 \\ -1 & 1 & 2 \end{bmatrix} \xrightarrow[R_3+1 \cdot R_1]{R_2-2R_1} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 3 & 1 \end{bmatrix} \xrightarrow{R_3+3R_2} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 4 \end{bmatrix} = U.$$

U is an echelon form (REF), not the RREF. The factors used were -2 (row 2 from row 1), $+1$ (row 3 from row 1) and $+3$ (row 3 from row 2). Their negatives go into L , at the colour-coded positions in the notes:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix}, \quad LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & -1 \\ -1 & 1 & 2 \end{bmatrix} = A. \checkmark$$

So L is a record of the row reduction from A to U .

Added example: using this LU to solve $Ax = b$

Take $b = (2, 4, 2)$ (here $P = I$). Step (2), $Ly = b$ by forward substitution: $y_1 = 2$; $2y_1 + y_2 = 4 \Rightarrow y_2 = 0$; $-y_1 - 3y_2 + y_3 = 2 \Rightarrow y_3 = 4$. Step (1), $Ux = y$ by back substitution: $4x_3 = 4 \Rightarrow x_3 = 1$; $-x_2 + x_3 = 0 \Rightarrow x_2 = 1$; $x_1 + 2x_2 - x_3 = 2 \Rightarrow x_1 = 1$. Check: $A(1, 1, 1)^T = (1 + 2 - 1, 2 + 3 - 1, -1 + 1 + 2)^T = (2, 4, 2)^T$. $\checkmark$

Why does this algorithm work? (p. 83)

The notes' argument. If A reduces to the echelon form U using only row replacements, then

$$E_k \cdots E_1 A = U,$$

where $E_1, \dots, E_k$ are the elementary matrices of the row replacements (9/10 notes, pp. 70–72). Each E_i is lower triangular, because each step adds a multiple of an *upper* row to a row *below* it, so the single off-diagonal entry of E_i sits below the diagonal. Multiplying by the inverses (each E_i is invertible, 9/10 notes, p. 71):

$$A = \underbrace{E_1^{-1} \cdots E_k^{-1}}_L U.$$

Each E_i^{-1} is also lower triangular (next box), and so is their product (Section 3). Each E_i^{-1} has 1s on the diagonal, so L does too.

Why the entries just drop into place: for the example, $E_1: R_2 - 2R_1$, $E_2: R_3 + R_1$, $E_3: R_3 + 3R_2$, and

$$E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix} = L. \checkmark$$

Each E_i^{-1} undoes its step, so it carries the factor with the opposite sign; in this order the products do not mix the entries.

Ex (p. 83): why is E_i^{-1} lower triangular? Because it is the opposite row replacement

Take $E_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ (row 3 $+= 2 \cdot$ row 1). Run the inverse algorithm $[E_i \ I] \rightarrow [I \ E_i^{-1}]$ (9/10 notes, p. 66):

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{R_3 - 2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 & 1 \end{array} \right], \quad E_i^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

One replacement step suffices, and it is “row 3 $-= 2 \cdot$ row 1”: the opposite operation, again adding a multiple of an upper row to a lower row. So E_i^{-1} is lower triangular. Check: $E_i E_i^{-1}$: row 3 is $(2, 0, 1) \cdot$ columns $= (2 - 2, 0, 1) = (0, 0, 1)$, so $E_i E_i^{-1} = I_3$. $\checkmark$

L , not L^{-1} , is the clean record. The product of the elementary matrices themselves is messier: for the example, $E_3 E_2 E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -5 & 3 & 1 \end{bmatrix}$ (the -5 mixes two steps). Its inverse L is the matrix whose entries are exactly the multipliers. *Fine print:* if a zero shows up in a pivot position, a row swap is needed and the simplifying assumption fails; that is what P in $A = PLU$ is for.

6 Subspaces of $\mathbb{R}^n$ (pp. 84–87)

A subspace is a set of vectors closed under addition and scaling; null spaces and spans are the main examples, and every subspace contains 0.

Definition (subspace, p. 84; Lay §2.8)

A set S of vectors in $\mathbb{R}^n$ is a **subspace** if it is closed under addition and scaling, that is,

1. u in S , c in $\mathbb{R} \Rightarrow cu$ in S ;
2. u, v in $S \Rightarrow u + v$ in S .

Theorem (p. 84)

S is a subspace $\Leftrightarrow$ for any $c_1, \dots, c_k$ in $\mathbb{R}$ and $u_1, \dots, u_k$ in S , the combination $c_1u_1 + \dots + c_ku_k$ is in S .

Proof of the theorem (not written in the notes)

$(\Leftarrow)$ Take $k = 1$ to get rule 1, and $k = 2$ with $c_1 = c_2 = 1$ to get rule 2. $(\Rightarrow)$ Each c_iu_i is in S by rule 1. Then $c_1u_1 + c_2u_2$ is in S by rule 2, then adding c_3u_3 keeps it in S , and so on, one term at a time (induction on k).

Example (p. 84): the solution space of $Ax = 0$ is a subspace (the “kernel”)

Rule 2: if $Ax = 0 = Ay$, then $A(x + y) = Ax + Ay = 0 + 0 = 0$.

$A(cx) \stackrel{?}{=} 0$: yes

$A(cx) = c(Ax) = c0 = 0$, so rule 1 holds too. Both steps use only the linearity of matrix multiplication (9/8 notes, p. 59).

Fill-in: for $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, $\text{Nul } A$ is the line $\text{span}\{(1, -2, 1)\}$ in $\mathbb{R}^3$

The notes draw $\text{Nul } A$ as a line through the origin. Row reduce: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \xrightarrow{R_2 - 4R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$, so $x_1 = x_3$, $x_2 = -2x_3$, x_3 free: $x = x_3(1, -2, 1)$. Check: $1 - 4 + 3 = 0$ and $4 - 10 + 6 = 0$. ✓

Definitions (p. 84)

$\text{Nul } A = \{x \text{ in } \mathbb{R}^n : Ax = 0\}$, the **null space** of the $m \times n$ matrix A . If $T(x) = Ax$, then the **kernel** of T is $\text{Ker } T = \text{Nul } A$.

Example (p. 85): a span is a subspace

If $v_1, \dots, v_k$ are in $\mathbb{R}^n$, then $\text{span}\{v_1, \dots, v_k\}$ is a subspace. The notes check both rules:

$$\begin{aligned} c(c_1v_1 + \dots + c_kv_k) &= (cc_1)v_1 + \dots + (cc_k)v_k, \\ (c_1v_1 + \dots + c_kv_k) + (d_1v_1 + \dots + d_kv_k) &= (c_1 + d_1)v_1 + \dots + (c_k + d_k)v_k. \end{aligned}$$

Both right sides are again combinations of $v_1, \dots, v_k$, so they lie in the span.

If $A = [a_1 \ \dots \ a_n]$ is $m \times n$, $\text{Col } A = \text{span}\{a_1, \dots, a_n\}$ is a subspace of $\mathbb{R}^m$

The **column space** $\text{Col } A$ is the span of the columns. Each column has m entries (one per row), so it lives in $\mathbb{R}^m$, and by the previous example the span is a subspace of $\mathbb{R}^m$. (Contrast: $\text{Nul } A$ lives in $\mathbb{R}^n$, because x has n entries.) Also $\text{Col } A = \{Ax : x \in \mathbb{R}^n\}$, the image of T (9/8 notes, p. 49).

Example (p. 85): any subspace contains 0, the zero vector

Scale by the number 0: if v is in S , then $0v = 0$ is in S by rule 1.

Fine print. This argument needs at least one v in S . The empty set satisfies rules 1 and 2 trivially, so Lay's definition (§2.8) also requires $0 \in S$, which rules it out. Practical test: *if $0 \notin S$, then S is not a subspace*. The converse fails (see the axes example below).

Ex (p. 86): for $b \neq 0$, is $\{x \in \mathbb{R}^n : Ax = b\}$ a subspace? No

It does not contain 0: $A0 = 0 \neq b$ (the notes' " $A0 \stackrel{!!}{=} b \neq 0$ "). By the previous example it is not a subspace. (If $Ax = b$ has solutions, they form a shifted copy of $\text{Nul } A$, Lay §1.5, which misses the origin.)

Ex (p. 86): is $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 = 0\}$ a subspace? No

S is the union of the two axes (the notes' picture). It contains 0 and is closed under scaling, but not under addition: $(1, 0)$ and $(0, 1)$ are in S , while $(1, 0) + (0, 1) = (1, 1)$ has $1 \cdot 1 = 1 \neq 0$. These are the two green arrows in the notes.

Ex (p. 86): is the unit square a subspace? No

Take the unit square $\{(x_1, x_2) : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$. The vector $(1, 1)$ is in it, but $2(1, 1) = (2, 2)$ is not (the long arrow in the notes leaving the square). So rule 1 fails. (Rule 1 also fails for $c = -1$.)

Ex (p. 87): is a line in $\mathbb{R}^2$ a subspace? Yes iff it passes through 0

Through 0: the line is $\{tv : t \in \mathbb{R}\} = \text{span}\{v\}$ for a nonzero direction v , and spans are subspaces. *Not through 0:* it misses 0, so it is not a subspace. Example: the line $x_2 = 1$ contains $(0, 1)$ but not $2(0, 1) = (0, 2)$.

Ex (p. 87): is $\{0\}$ a subspace of $\mathbb{R}^n$? Yes

$c0 = 0$ and $0 + 0 = 0$, so both rules hold. It is the smallest subspace ($\{0\} = \text{Nul } I_n = \text{span}\{0\}$).

7 Bases, and a basis of $\text{Nul } A$ (pp. 87–89)

A basis is an independent spanning set; solving $Ax = 0$ in parametric form produces a basis of $\text{Nul } A$.

Definition (basis, p. 87; Lay §2.8)

A **basis** of a subspace S is a linearly independent set of vectors in S that spans S .

Ex (p. 88): the standard basis $\{e_1, \dots, e_n\}$ is a basis of $\mathbb{R}^n$. Why?

Independent: $c_1e_1 + \dots + c_ne_n = (c_1, \dots, c_n)$, so if it is 0 then $c_i = 0$ for each i . *Spanning:* every $x = (x_1, \dots, x_n)$ equals $x_1e_1 + \dots + x_ne_n$. (Equivalently: the matrix $[e_1 \ \dots \ e_n] = I_n$ is invertible, so by the IMT (e) and (h) its columns are independent and span $\mathbb{R}^n$.)

Example (pp. 88–89): find a basis of $\text{Nul } A$

$$A = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 3 & 0 & 1 \end{bmatrix}. \text{ Process: solve } Ax = 0.$$

Basis of $\text{Nul } A$: $\{(-3, 2, 1, 0), (1, -1, 0, 1)\}$ (the notes' RREF has a sign slip; see Corrections)

Row reduce:

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

(The notes have -5 and 3 in the first row; the correct entries are $-1 - 2(-2) = 3$ and $1 - 2(1) = -1$.) Columns 1, 2 are pivotal and x_3, x_4 are free:

$$x_1 = -3x_3 + x_4, \quad x_2 = 2x_3 - x_4,$$

so

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

Why these two vectors are a basis: they span $\text{Nul } A$ by the display (every solution is such a combination). They are independent because of the free-variable slots: in entries 3 and 4 they look like $(1, 0)$ and $(0, 1)$, so $x_3v_1 + x_4v_2 = 0$ forces $x_3 = x_4 = 0$.

Check: $A(-3, 2, 1, 0)^T = (-3 + 4 - 1, -6 + 6 + 0)^T = (0, 0)^T$ and $A(1, -1, 0, 1)^T = (1 - 2 + 1, 2 - 3 + 1)^T = (0, 0)^T$. ✓ (The notes' vectors $(5, 2, 1, 0)$ and $(-3, -1, 0, 1)$ give $A(5, 2, 1, 0)^T = (8, 16)^T \neq 0$, so they are not in $\text{Nul } A$.)

$T(x) = Ax$, $T : \mathbb{R}^4 \rightarrow \mathbb{R}^2$, is onto. Why? And $\text{Nul } A = T^{-1}(0)$

Onto: the echelon form has a pivot in each of its 2 rows, so $Ax = b$ is solvable for every $b \in \mathbb{R}^2$ (9/8 notes, p. 49; Lay §1.4, Theorem 4). It is not one-to-one: columns 3 and 4 are not pivotal, so $\text{Nul } A \neq \{0\}$. *Meaning of $T^{-1}(0)$:* the set of all inputs that T sends to 0, which is $\{x : Ax = 0\} = \text{Nul } A = \text{Ker } T$. Here T^{-1} means preimage; T itself has no inverse function (it is not one-to-one).

Looking ahead (9/17). The general theorem behind this example ("the usual algorithm for solving $Ax = 0$ gives a basis of $\text{Nul } A$ ", proof by this example) and the analogous fact for

Col A (pivot columns form a basis) open the next lecture (p. 90), together with dimension and rank.

Summary

Fact	Formula / statement
IMT ($n \times n$)	invertible $\Leftrightarrow n$ pivots $\Leftrightarrow Ax = 0$ only trivially $\Leftrightarrow T$ one-to-one $\Leftrightarrow T$ onto $\Leftrightarrow CA = I \Leftrightarrow AD = I \Leftrightarrow A^T$ invertible
Lower $\times$ lower	lower triangular (same for upper)
Permuted LU	$A = PLU$; L unit lower triangular, U an echelon form
Solving with LU	$Ly = P^{-1}b$ first, then $Ux = y$; $O(n^2)$ each
Building L (no swaps)	step $R_i += cR_j$ puts $-c$ at position (i, j)
Subspace	closed under $+$ and scaling; contains 0
Examples of subspaces	$\text{Nul } A \subseteq \mathbb{R}^n$, $\text{Col } A \subseteq \mathbb{R}^m$, $\text{span}\{v_1, \dots, v_k\}$, $\{0\}$, lines through 0
Basis	independent and spanning; parametric solution of $Ax = 0$ gives a basis of $\text{Nul } A$

Corrections to the handwritten notes

Page	As written	Should be	Why
75	$b \Leftrightarrow c \Leftrightarrow d \Leftrightarrow c$ (IMT diagram)	$b \Leftrightarrow c \Leftrightarrow d \Leftrightarrow e$	(e) is otherwise never linked; $d \Leftrightarrow e$ is the next line of the notes
78	$P(x_1, x_2, x_n)$	$P(x_1, x_2, x_3)$	$n = 3$ here (same value)
80	$x_2 = \frac{1-\frac{1}{3}}{2} = \frac{1}{3}$	$x_1 = \frac{1-\frac{1}{3}}{2} = \frac{1}{3}$	solving $2x_1 + x_2 = 1$ for x_1
88	RREF row 1: $[1 \ 0 \ -5 \ 3]$	$[1 \ 0 \ 3 \ -1]$	$R_1 - 2R_2: -1 - 2(-2) = 3$, $1 - 2(1) = -1$
88	$x_1 = 5x_3 - 3x_4$	$x_1 = -3x_3 + x_4$	from the corrected row 1
89	basis $(5, 2, 1, 0)$, $(-3, -1, 0, 1)$	$(-3, 2, 1, 0)$, $(1, -1, 0, 1)$	$A(5, 2, 1, 0)^T = (8, 16)^T \neq 0$