Lecture 14: Counting Solutions and the Matrix Inverse

Elementary matrices, invertibility (§3.2–3.3)

Differential Equations & Linear Algebra

1 Recap: REF and RREF

Last time: linear systems, augmented matrices, and elementary row operations,

$$\mathbf{A}\vec{x} = \vec{b} \quad \longrightarrow \quad [\,\mathbf{A} \mid \vec{b}\,].$$

For $[\,\mathbf{A} \mid \vec{b}\,]$ to be in REF and/or RREF, it must satisfy:

(i) Zero rows (if any) are on the bottom.
(ii) Leftmost *nonzero* entry of each nonzero row (the "pivot") is 1.
(iii) Each pivot is to the right of the pivots above it.

(i)–(iii): **REF**

(iv) Each pivot is the only nonzero entry in its column.

(i)–(iv): **RREF**

1.1 Examples

Pivots are circled.

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$ breaks (ii) and (iv)

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$ RREF ✓

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$ breaks (iv) $\Rightarrow$ REF, not RREF

$$\begin{bmatrix} 1 & 3 & 0 & -1 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$ RREF ✓

2 Reading off the number of solutions

Once a system is in REF or RREF, we can *immediately* identify whether 0, 1, or ∞ many solutions exist — scenarios we describe with the following language:

- If no solutions exist, the system is **inconsistent**
- If 1 or more solutions exist, the system is **consistent**
- If ∞ solutions exist, the system is consistent but **underdetermined**

We identify these from REF/RREF as follows:

- If any row looks like $\left[\begin{array}{ccc|c} 0 & \cdots & 0 & k \end{array}\right]$ with $k \neq 0$, then there are **no solutions**: the system is **inconsistent**.
- Otherwise the system is **consistent**: there is at least one solution.

For consistent systems:

- If every column in $\mathbf{A}$ (not in $[\mathbf{A} \mid \vec{b}]$) is a pivot column, there is exactly **one unique solution**.
  - Recall that the **rank** of a matrix, denoted $\operatorname{rank}\mathbf{A}$, is defined as the number of pivot columns; so this is saying that the rank of $\mathbf{A}$ must equal the number of independent variables in the system to have a unique solution.
- If there are one or more non-pivot columns in $\mathbf{A}$ (not in $[\mathbf{A} \mid \vec{b}]$), there are ∞ **many solutions**: the system is **underdetermined**.

See Ex. 5.1 at the end of the Lecture 13 notes; also §3.2, Examples 6–9.

3 Row reduction as matrix multiplication

Row reductions are the matrix equivalent of manipulating individual equations in high-school algebra.

High-school algebra:

$$\begin{aligned} & x^2 + y^2 = 1 && \text{subtract } x^2 \text{ from both sides} \\ \Leftrightarrow \quad & (x^2 + y^2) - x^2 = 1 - x^2 && \\ \Leftrightarrow \quad & y^2 = 1 - x^2 && \text{take } \sqrt{\cdot} \text{ of both sides} \\ \Leftrightarrow \quad & y = \pm\sqrt{1 - x^2} && \end{aligned}$$

Central logic: if $a = b$ is a true statement, then so is $a \cdot c = b \cdot c$, and $a^2 = b^2$, etc. We just need to do the *same* manipulation to both sides.

Here: if the statement $\mathbf{A}\vec{x} = \vec{b}$ is true, then we can manipulate both sides and write down other true statements (kind of like *equivalent systems*, discussed last time). For instance,

$$\mathbf{A}\vec{x} = \vec{b} \quad \Longrightarrow \quad \mathbf{B}\mathbf{A}\vec{x} = \mathbf{B}\vec{b}.$$

(Note: **order matters!** Usually $\mathbf{AB} \neq \mathbf{BA}$.)

Example 3.1. Suppose

$$\overbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}^{\mathbf{A}} \overbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}^{\vec{x}} = \overbrace{\begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}}^{\vec{b}}.$$

(Note that we can see immediately, by doing the matrix multiplication on the left-hand side of the "=", that $2x = 2$, $z = 3$, and $y = 2$.)

We can multiply both sides from the left by

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \text{to get} \qquad \mathbf{E}_1\mathbf{A}\vec{x} = \mathbf{E}_1\vec{b}:$$

$$\Longleftrightarrow \quad \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{\text{do this multiplication}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}}_{\text{\& do this multiplication}}$$

$$\Longleftrightarrow \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

Now multiply on the left by $\mathbf{E}_2 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$:

$$\Longleftrightarrow \quad \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$

$$\Longleftrightarrow \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \text{— hey, that's RREF!}$$

3.1 Elementary matrices

Elementary row operations R_j on $[\mathbf{A} \mid \vec{b}] \Longleftrightarrow$ left-multiplication of $\mathbf{A}\vec{x} = \vec{b}$ by **elementary matrices** $\mathbf{E}_j$, where the elementary matrix $\mathbf{E}_j$ is the result of doing R_j to $\mathbb{I}$.

In Example 3.1, $\mathbf{E}_1$ is $\mathbb{I}$ with rows 2 and 3 exchanged, and $\mathbf{E}_2$ is $\mathbb{I}$ with row 1 multiplied by $\frac{1}{2}$.

Now consider $\mathbf{A}\vec{x} = \vec{b}$ again, and left-multiply by $\mathbb{I}$:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{\text{do this multiplication}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}}_{\text{leave this}}$$

$$\iff \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}.$$

As before, left-multiply by

$$\mathbf{E}_2\mathbf{E}_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

So

Note that, for any square matrix $\mathbf{B}$, we have $\mathbf{B}\mathbb{I} = \mathbb{I}\mathbf{B} = \mathbf{B}$.

$$\begin{aligned} & \mathbf{A}\vec{x} = \vec{b} \\ \iff \quad & \mathbf{A}\vec{x} = \mathbb{I}\vec{b} \\ \iff \quad & \underbrace{\mathbf{E}_2\mathbf{E}_1\mathbf{A}}_{\substack{\text{already showed} \\ \text{this} = \mathbb{I}}}\vec{x} = \underbrace{\mathbf{E}_2\mathbf{E}_1\mathbb{I}}_{\substack{\text{use } \mathbf{E}_2\mathbf{E}_1\mathbb{I} = \mathbf{E}_2\mathbf{E}_1 \\ \text{(property of } \mathbb{I})}}\vec{b} \\ \iff \quad & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\vec{x} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\vec{b} \\ \iff \quad & \mathbb{I}\vec{x} = \mathbf{E}_2\mathbf{E}_1\vec{b} \\ \iff \quad & \vec{x} = \mathbf{E}_2\mathbf{E}_1\vec{b}. \end{aligned}$$

To recap, we have

$$\mathbf{A}\vec{x} = \vec{b} \quad \iff \quad \vec{x} = \mathbf{E}_2\mathbf{E}_1\vec{b},$$

and we also showed that $\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbb{I}$.
This is kind of like how some nonzero scalar a multiplied by $(a)^{-1} = 1/a$ always produces 1 (i.e., $a/a = 1$). There's no such thing as "dividing by $\mathbf{A}$" for a matrix $\mathbf{A}$ the way you can divide things by scalars, but this is the next best thing.

4 The matrix inverse

Recall: matrices represent linear operators; they're a special kind of function whose input and output are both vectors. In the example above,

$$\left.\begin{array}{l}\mathbf{A} \text{ maps } \vec{x} \text{ to } \vec{b} \\ \mathbf{E}_2\mathbf{E}_1 \text{ maps } \vec{b} \text{ to } \vec{x}\end{array}\right\} \longrightarrow \begin{array}{l}\text{like how, if } y = x^2, \text{ then } \sqrt{y} = x\text{:} \\ (\cdot)^2 \text{ maps } x \mapsto y \text{ and } \sqrt{\cdot} \text{ maps } y \mapsto x \\ \text{(for } x \geq 0 \text{ anyway).}\end{array}$$

So $\mathbf{E}_2\mathbf{E}_1$ undoes $\mathbf{A}$ like how $\sqrt{\cdot}$ undoes $(\cdot)^2$. Or we could say $\mathbf{E}_2\mathbf{E}_1$ does the *inverse* of $\mathbf{A} \implies$ let's call it $\mathbf{A}^{-1}$.

4.1 Definition and properties

Matrix inverse. For an $n \times n$ matrix $\mathbf{A}$, if

$$\underbrace{\operatorname{rank}\mathbf{A}}_{\#\text{ of pivot columns}} = n,$$

then $\mathbf{A}$ is **invertible**, and there exists $\mathbf{A}^{-1}$ with the following properties:

- $\mathbf{A}\vec{x} = \vec{b} \iff \vec{x} = \mathbf{A}^{-1}\vec{b}$.
- $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbb{I}$.
- If $R_1, R_2, \ldots, R_k$ are the row operations done to $\mathbf{A}$ to yield $\mathbb{I}$ (its RREF), then doing $R_1, R_2, \ldots, R_k$ to $\mathbb{I}$ yields $\mathbf{A}^{-1}$.
  - Hence the **augmented matrix method** for computing $\mathbf{A}^{-1}$ (See p. 149 of the book.):

$$\left[\mathbf{A} \mid \mathbb{I}\right] \xrightarrow{\text{do } R_1,\ldots,R_k} \left[\mathbb{I} \mid \mathbf{A}^{-1}\right].$$

More properties of inverses:

- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
- $(\mathbf{A}^{-1})^{\mathsf{T}} = (\mathbf{A}^{\mathsf{T}})^{-1}$.
- If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ and invertible, then

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \qquad (\underline{\textbf{not}}\ \mathbf{A}^{-1}\mathbf{B}^{-1}).$$

Note. If $\mathbf{A}$'s RREF is not $\mathbb{I}_n$ (i.e., if $\operatorname{rank}\mathbf{A} < n$), then $\mathbf{A}$ is **not invertible**.

4.2 The "Fundamental" Theorem of Invertible Matrices

$n \times n$ matrices have a long list of properties that are true if and only if they are invertible; we've now introduced enough concepts to start listing some of these properties, with more to come later.

Theorem

If $\mathbf{A}$ is $n \times n$, then the following are equivalent:

(a) $\mathbf{A}$ has n pivot columns.

(b) $\mathbf{A}$'s RREF is $\mathbb{I}_n$.

(c) $\mathbf{A}$ is invertible, i.e., $\mathbf{A}^{-1}$ exists.

(d) $\mathbf{A}^{-1}$ is invertible.

(e) $\operatorname{rank} \mathbf{A} = n$.

(f) $\mathbf{A}^{\mathsf{T}}$ is invertible.

(g) For every $\vec{b} \in \mathbb{R}^n$, the system $\mathbf{A}\vec{x} = \vec{b}$ is consistent and $\vec{x}$ is unique.

(h) $\mathbf{A}\vec{x} = \vec{0} \iff \vec{x} = \vec{0}$.

(To be continued...)

(For an example of a case where condition (g) above is not met, go back to the underdetermined system example at the end of last lecture, or Examples 6 – 7 in §3.2 of the book. This is especially important for later on when we apply these ideas to differential equations.)