

Lecture 10: Matrix Inversion

As we pointed out in Lecture 9, existence of multiplicative inverses of matrices is a messy issue for non-square matrices, and even square matrices may or may not have one. Let's begin our study of inverses in earnest with 2×2 matrices.

$$\text{Ex 1/} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & \cancel{ab-ab} \\ \cancel{cd-cd} & ad-bc \end{pmatrix} = (ad-bc) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

So if $ad-bc \neq 0$,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{I}_2,$$

making $\frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ "an inverse" to $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ on the right. //

Some Questions: (1) Does it work on the left too?

(2) Is it unique?

(3) Is there an algorithm for producing it (not just for 2×2)?

(As we'll see, the answer is YES for all three.)

Inversion Algorithm $A = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{v}_1 & \dots & \vec{v}_n \\ \downarrow & & \downarrow \end{pmatrix}$ $n \times n$ matrix

Want an $n \times n$ $B = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{w}_1 & \dots & \vec{w}_n \\ \downarrow & & \downarrow \end{pmatrix}$ with $AB = \mathbb{I}_n$, i.e.

$$\begin{pmatrix} \uparrow & & \uparrow \\ A\vec{w}_1 & \dots & A\vec{w}_n \\ \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{e}_1 & \dots & \vec{e}_n \\ \downarrow & & \downarrow \end{pmatrix},$$

which is equivalent to solving n systems

$$(*) \quad A\vec{w}_1 = \vec{e}_1, \dots, A\vec{w}_n = \vec{e}_n$$

for $\vec{w}_1, \dots, \vec{w}_n$. Since $A\vec{w}_k = A \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix} = b_{1k}\vec{v}_1 + \dots + b_{nk}\vec{v}_n$, we must have that

$$\text{span}\{\vec{v}_1, \dots, \vec{v}_n\} \text{ contains } \vec{e}_1, \dots, \vec{e}_n$$

$$\Rightarrow \text{span}\{\vec{v}_1, \dots, \vec{v}_n\} = \mathbb{R}^n$$

$$\Rightarrow \text{ref}(A) \text{ has a leading 1 in each row}$$

$$\Rightarrow \text{ref}(A) = \mathbb{I}_n.$$

Conversely, if $\text{ref}(A) = \mathbb{I}_n$, then (for each k)

$$\text{ref}[A | \vec{e}_k] = [\mathbb{I}_n | \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}]$$

and then $\begin{cases} x_1 = c_1 \\ \vdots \\ x_n = c_n \end{cases}$ solves the system $A\vec{x} = \vec{e}_k$ — that is,

$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ is the desired $\vec{w}_k$. To solve all n systems

simultaneously, take

$$\text{rref}[A \mid \mathbb{I}_n] = [\mathbb{I}_n \mid ?]$$

and then the "?" will give you B.

Ex 2/ Given $A = \begin{pmatrix} 1 & -1 & 0 \\ -2 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$, we take rref of

$$\begin{aligned} \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ -2 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right] \\ &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right]. \end{aligned}$$

So $B = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}$ satisfies $AB = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{I}_3$. //

Ex 3/ $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ again. Write $\Delta = ad - bc$, assumed $\Delta \neq 0$.

The row-reduction breaks into 2 cases: $a = 0$ & $a \neq 0$. I'll do the $a \neq 0$ case:

$$\begin{aligned} [A \mid \mathbb{I}_2] &\rightarrow \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ c & d & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & d - \frac{cb}{a} & -\frac{c}{a} & 1 \end{array} \right] \\ &\rightarrow \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & 1 & -\frac{c}{\Delta} & \frac{a}{\Delta} \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & \frac{bc}{a\Delta} + \frac{1}{a} & -\frac{b}{\Delta} \\ 0 & 1 & -\frac{c}{\Delta} & \frac{a}{\Delta} \end{array} \right], \text{ and } \frac{bc}{a\Delta} + \frac{1}{a} = \frac{bc + \Delta}{a\Delta} \\ &= \frac{bc + ad - bc}{a\Delta} = \frac{ad}{a\Delta} = \frac{d}{\Delta} \\ \Rightarrow B &= \begin{pmatrix} \frac{d}{\Delta} & -\frac{b}{\Delta} \\ -\frac{c}{\Delta} & \frac{a}{\Delta} \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \end{aligned}$$

REF RRevisited

Recall the three types of row operations on a matrix:

- Replace
 - Swap
 - Scale
- I claim that these can be achieved by left-multiplication by certain special matrices called elementary matrices.

Ex 4/

REPLACE $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 0 \\ -2 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$

SWAP $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

SCALE $\begin{pmatrix} 1/a & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & b/a \\ c & d \end{pmatrix}$

Elementary Matrices

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & a & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \cdot \begin{pmatrix} \leftarrow & \vec{r}_1 & \rightarrow \\ \leftarrow & \vec{r}_2 & \rightarrow \\ \leftarrow & \vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_j + a\vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_{j+1} & \rightarrow \\ \leftarrow & \vec{r}_n & \rightarrow \end{pmatrix} = \begin{pmatrix} \leftarrow & \vec{r}_1 & \rightarrow \\ \leftarrow & \vec{r}_2 & \rightarrow \\ \leftarrow & \vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_j + a\vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_{j+1} & \rightarrow \\ \leftarrow & \vec{r}_n & \rightarrow \end{pmatrix}$$

row j
column i

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & & 1 \\ & & & \ddots & \\ & & 1 & & 0 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \cdot \begin{pmatrix} \leftarrow & \vec{r}_1 & \rightarrow \\ \leftarrow & \vec{r}_2 & \rightarrow \\ \leftarrow & \vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_j & \rightarrow \\ \leftarrow & \vec{r}_k & \rightarrow \end{pmatrix} = \begin{pmatrix} \leftarrow & \vec{r}_1 & \rightarrow \\ \leftarrow & \vec{r}_2 & \rightarrow \\ \leftarrow & \vec{r}_j & \rightarrow \\ \leftarrow & \vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_k & \rightarrow \end{pmatrix}$$

row i
row j
column i
column j

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & a & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \cdot \begin{pmatrix} \leftarrow & \vec{r}_1 & \rightarrow \\ \leftarrow & \vec{r}_2 & \rightarrow \\ \leftarrow & \vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_j & \rightarrow \end{pmatrix} = \begin{pmatrix} \leftarrow & \vec{r}_1 & \rightarrow \\ \leftarrow & \vec{r}_2 & \rightarrow \\ \leftarrow & a\vec{r}_i & \rightarrow \\ \leftarrow & \vec{r}_j & \rightarrow \end{pmatrix}$$

row i
column i

The elementary matrices are the 3 kinds of matrices on the left, producing Replace, Swap, & Scale operations respectively. They are $n \times n$.

Upshot: The result of any sequence of row operations on any matrix A can be expressed as

$$E_N \cdots E_1 \cdot A$$

where the E_i are elementary matrices.

Hence, if $\text{rref}(A) = I_n$, then

$$E \cdot A = I_n \quad (\text{where } E = \text{product of elementary matrices})$$

and

$$\text{rref}[A | I_n] = E \cdot [A | I_n] = [EA | EI_n] = [I_n | E],$$

giving another perspective on why the inversion algorithm works.

Conversely, if $CA = I_n$ (for some $n \times n$ matrix C), then

$$A\vec{x} = \vec{0} \implies \vec{x} = CA\vec{x} = C\vec{0} = \vec{0}$$

(i.e. $A\vec{x} = \vec{0}$ has only the trivial solution) and A has no non-pivot columns. Hence $\text{rref}(A) = I_n$.

Moreover, if $CA = I_n = AB$, then

$$B = I_n B = CAB = CI_n = C.$$

This brings us to ...

Theorem/Definition: An $n \times n$ matrix A is invertible if (and only if) one of the equivalent statements

$$\begin{cases} \text{(i)} & \text{rref}(A) = I_n \\ \text{(ii)} & \text{there's a matrix } B \text{ with } AB = I_n \\ \text{(iii)} & \text{" " " } C \text{ " } CA = I_n \end{cases}$$

holds. The inverse of A is then $A^{-1} := B = C$.

Comments:

① Since elementary row operations transform A to the identity and are reversible, they also transform I_n to A . More precisely,

$$A = E_1^{-1} E_2^{-1} \cdots E_N^{-1},$$

so any invertible matrix is a product of elementary matrices.

② $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible $\iff ad - bc \neq 0$.

(We already know " $\Leftarrow$ " by Example 1. If $ad - bc = 0$, then $ad = bc \Rightarrow$ rows are proportional $\Rightarrow \text{rref}(A) = \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} \Rightarrow A$ not invertible.)

Properties of the Inverse

Let $A, B =$ invertible $n \times n$ matrices. Then

- $A^{-1} \cdot A = \mathbb{I}_n = A \cdot A^{-1}$
- $(A^{-1})^{-1} = A$
- $(A^T)^{-1} = (A^{-1})^T$ (b/c $A^T(A^{-1})^T = (A^{-1}A)^T = \mathbb{I}_n^T = \mathbb{I}_n$)
- $(AB)^{-1} = B^{-1}A^{-1}$ (b/c $(AB) \cdot (B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\mathbb{I}_n A^{-1} = AA^{-1} = \mathbb{I}_n$)

Using the inverse to solve inhomogeneous systems

Ex / Solve $\underbrace{\begin{pmatrix} 5 & -9 \\ -4 & 7 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} 1 \\ -1 \end{pmatrix}}_{\vec{b}}$

Multiply both sides of $A\vec{x} = \vec{b}$ on the left by A^{-1} :

$$\vec{x} = A^{-1}A\vec{x} = A^{-1}\vec{b}$$

$$= \underbrace{\frac{1}{5 \cdot 7 - (-4)(-9)} \begin{pmatrix} 7 & 9 \\ 4 & 5 \end{pmatrix}}_{A^{-1}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= - \begin{pmatrix} 7 & 9 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = - \begin{pmatrix} -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} //$$

... goes a bit faster than row-reducing the augmented matrix!