MATH 323 Linear Algebra

Lecture 5: Inverse matrix (continued).

# **Identity matrix**

Definition. The **identity matrix** (or **unit matrix**) is a diagonal matrix with all diagonal entries equal to 1. The  $n \times n$  identity matrix is denoted  $I_n$  or simply I.

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
  
In general,  $I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$ 

**Theorem.** Let A be an arbitrary  $m \times n$  matrix. Then  $I_m A = A I_n = A$ .

#### **Inverse matrix**

Definition. Let A be an  $n \times n$  matrix. The **inverse** of A is an  $n \times n$  matrix, denoted  $A^{-1}$ , such that

$$AA^{-1} = A^{-1}A = I.$$

If  $A^{-1}$  exists then the matrix A is called **invertible**. Otherwise A is called **singular**.

Let A and B be  $n \times n$  matrices. If A is invertible then we can **divide** B by A:

left division:  $A^{-1}B$ , right division:  $BA^{-1}$ .

*Remark.* There is no notation for the matrix division and the notion is not really used.

Basic properties of inverse matrices

• If  $B = A^{-1}$  then  $A = B^{-1}$ . In other words, if A is invertible, so is  $A^{-1}$ , and  $A = (A^{-1})^{-1}$ .

• The inverse matrix (if it exists) is unique. Moreover, if AB = CA = I for some  $n \times n$  matrices B and C, then  $B = C = A^{-1}$ .

Indeed, B = IB = (CA)B = C(AB) = CI = C.

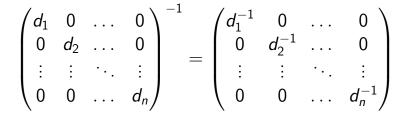
• If  $n \times n$  matrices A and B are invertible, so is AB, and  $(AB)^{-1} = B^{-1}A^{-1}$ .

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I,$$
  
 $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$ 

• By induction,  $(A_1A_2...A_k)^{-1} = A_k^{-1}...A_2^{-1}A_1^{-1}$ .

### Inverting diagonal matrices

**Theorem** A diagonal matrix  $D = \text{diag}(d_1, \ldots, d_n)$ is invertible if and only if all diagonal entries are nonzero:  $d_i \neq 0$  for  $1 \leq i \leq n$ . If D is invertible then  $D^{-1} = \text{diag}(d_1^{-1}, \ldots, d_n^{-1})$ .



## **Inverting diagonal matrices**

**Theorem** A diagonal matrix  $D = \text{diag}(d_1, \ldots, d_n)$  is invertible if and only if all diagonal entries are nonzero:  $d_i \neq 0$  for  $1 \leq i \leq n$ .

If D is invertible then  $D^{-1} = \operatorname{diag}(d_1^{-1}, \ldots, d_n^{-1})$ .

*Proof:* If all  $d_i \neq 0$  then, clearly,  $\operatorname{diag}(d_1, \ldots, d_n) \operatorname{diag}(d_1^{-1}, \ldots, d_n^{-1}) = \operatorname{diag}(1, \ldots, 1) = I$ ,  $\operatorname{diag}(d_1^{-1}, \ldots, d_n^{-1}) \operatorname{diag}(d_1, \ldots, d_n) = \operatorname{diag}(1, \ldots, 1) = I$ . Now suppose that  $d_i = 0$  for some *i*. Then for any  $n \times n$  matrix *B* the *i*th row of the matrix *DB* is a

zero row. Hence  $DB \neq I$ . Therefore D is singular.

## Inverting 2×2 matrices

Definition. The **determinant** of a 2×2 matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  is det A = ad - bc.

**Theorem** A matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  is invertible if and only if det  $A \neq 0$ .

If det  $A \neq 0$  then  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$  **Theorem** A matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  is invertible if and only if det  $A \neq 0$ . If det  $A \neq 0$  then  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$ *Proof:* Let  $B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ . Then  $AB = BA = \begin{pmatrix} ad-bc & 0\\ 0 & ad-bc \end{pmatrix} = (ad-bc)I_2.$ 

In the case det  $A \neq 0$ , we have  $A^{-1} = (\det A)^{-1}B$ . In the case det A = 0, the matrix A is not invertible as otherwise  $AB = O \implies A^{-1}(AB) = A^{-1}O = O$  $\implies (A^{-1}A)B = O \implies I_2B = O \implies B = O$  $\implies A = O$ , but the zero matrix is singular.

# **Problem.** Solve a system $\begin{cases} 4x + 3y = 5, \\ 3x + 2y = -1. \end{cases}$

This system is equivalent to a matrix equation  $A\mathbf{x} = \mathbf{b}$ ,

where 
$$A = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix}$$
,  $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ ,  $\mathbf{b} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$ .

We have det  $A = -1 \neq 0$ . Hence A is invertible.

$$A\mathbf{x} = \mathbf{b} \implies A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b} \implies (A^{-1}A)\mathbf{x} = A^{-1}\mathbf{b}$$
$$\implies \mathbf{x} = A^{-1}\mathbf{b}.$$

Conversely,  $\mathbf{x} = A^{-1}\mathbf{b} \implies A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = \mathbf{b}$ .

Thus the system has a unique solution

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \frac{1}{-1} \begin{pmatrix} 2 & -3 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \begin{pmatrix} -13 \\ 19 \end{pmatrix}$$

System of *n* linear equations in *n* variables:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \iff A\mathbf{x} = \mathbf{b},$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

**Theorem** If the matrix A is invertible then the system has a unique solution, which is  $\mathbf{x} = A^{-1}\mathbf{b}$ .

## General results on inverse matrices

**Theorem 1** Given an  $n \times n$  matrix A, the following conditions are equivalent:

(i) A is invertible;

(ii)  $\mathbf{x} = \mathbf{0}$  is the only solution of the matrix equation  $A\mathbf{x} = \mathbf{0}$ ; (iii) for some *n*-dimensional column vector **b**, the matrix equation  $A\mathbf{x} = \mathbf{b}$  has a unique solution (which is  $\mathbf{x} = A^{-1}\mathbf{b}$ ); (iv) the matrix equation  $A\mathbf{x} = \mathbf{b}$  has a unique solution for any *n*-dimensional column vector **b**;

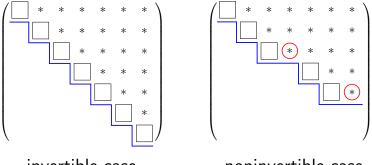
(v) the row echelon form of A has no zero rows;

(vi) the reduced row echelon form of A is the identity matrix.

**Theorem 2** Suppose that a sequence of elementary row operations converts a matrix *A* into the identity matrix.

Then the same sequence of operations converts the identity matrix into the inverse matrix  $A^{-1}$ .

### Row echelon form of a square matrix:

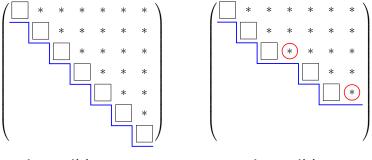


invertible case

noninvertible case

For any matrix in row echelon form, the number of columns with leading entries equals the number of rows with leading entries. For a square matrix, also the number of columns *without* leading entries (i.e., the number of free variables in a related system of linear equations) equals the number of rows *without* leading entries (i.e., zero rows).

### Row echelon form of a square matrix:



invertible case

noninvertible case

Hence the row echelon form of a square matrix A is either strict triangular or else it has a zero row. In the former case, the equation  $A\mathbf{x} = \mathbf{b}$  always has a unique solution. In the latter case,  $A\mathbf{x} = \mathbf{b}$  never has a unique solution. Also, in the former case the reduced row echelon form of A is I.

Example. 
$$A = \begin{pmatrix} 3 & -2 & 0 \\ 1 & 0 & 1 \\ -2 & 3 & 0 \end{pmatrix}$$
.

To check whether A is invertible, we convert it to row echelon form.

Interchange the 1st row with the 2nd row:

$$\begin{pmatrix} 1 & 0 & 1 \\ 3 & -2 & 0 \\ -2 & 3 & 0 \end{pmatrix}$$

Add -3 times the 1st row to the 2nd row:

$$egin{pmatrix} 1 & 0 & 1 \ 0 & -2 & -3 \ -2 & 3 & 0 \end{pmatrix}$$

Add 2 times the 1st row to the 3rd row:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & -3 \\ 0 & 3 & 2 \end{pmatrix}$$

# Multiply the 2nd row by -0.5:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1.5 \\ 0 & 3 & 2 \end{pmatrix}$$

Add -3 times the 2nd row to the 3rd row:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1.5 \\ 0 & 0 & -2.5 \end{pmatrix}$$

# Multiply the 3rd row by -0.4: $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1.5 \\ 0 & 0 & 1 \end{pmatrix}$

We already know that the matrix A is invertible. Let's proceed towards reduced row echelon form.

Add -1.5 times the 3rd row to the 2nd row:  $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ 

Add -1 times the 3rd row to the 1st row:  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  To obtain  $A^{-1}$ , we need to apply the following sequence of elementary row operations to the identity matrix:

- interchange the 1st row with the 2nd row,
- add -3 times the 1st row to the 2nd row,
- add 2 times the 1st row to the 3rd row,
- multiply the 2nd row by -0.5,
- add -3 times the 2nd row to the 3rd row,
- multiply the 3rd row by -0.4,
- add -1.5 times the 3rd row to the 2nd row,
- add -1 times the 3rd row to the 1st row.

A convenient way to compute the inverse matrix  $A^{-1}$  is to merge the matrices A and I into one  $3 \times 6$  matrix  $(A \mid I)$ , and apply elementary row operations to this new matrix.

$$A = \begin{pmatrix} 3 & -2 & 0 \\ 1 & 0 & 1 \\ -2 & 3 & 0 \end{pmatrix}, \qquad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$(A \mid I) = \begin{pmatrix} 3 & -2 & 0 \mid 1 & 0 & 0 \\ 1 & 0 & 1 \mid 0 & 1 & 0 \\ -2 & 3 & 0 \mid 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & -2 & 0 & | & 1 & 0 & 0 \\ 1 & 0 & 1 & | & 0 & 1 & 0 \\ -2 & 3 & 0 & | & 0 & 0 & 1 \end{pmatrix}$$

Interchange the 1st row with the 2nd row:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 3 & -2 & 0 & 1 & 0 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Add -3 times the 1st row to the 2nd row:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -2 & -3 & 1 & -3 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Add 2 times the 1st row to the 3rd row:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -2 & -3 & 1 & -3 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{pmatrix}$$

Multiply the 2nd row by -0.5:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1.5 & -0.5 & 1.5 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{pmatrix}$$

Add -3 times the 2nd row to the 3rd row:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1.5 & -0.5 & 1.5 & 0 \\ 0 & 0 & -2.5 & 1.5 & -2.5 & 1 \end{pmatrix}$$

Multiply the 3rd row by -0.4:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1.5 & -0.5 & 1.5 & 0 \\ 0 & 0 & 1 & -0.6 & 1 & -0.4 \end{pmatrix}$$

Add -1.5 times the 3rd row to the 2nd row:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0.4 & 0 & 0.6 \\ 0 & 0 & 1 & -0.6 & 1 & -0.4 \end{pmatrix}$$

Add -1 times the 3rd row to the 1st row:

$$egin{pmatrix} 1 & 0 & 0 & 0.6 & 0 & 0.4 \ 0 & 1 & 0 & 0.4 & 0 & 0.6 \ 0 & 0 & 1 & -0.6 & 1 & -0.4 \end{pmatrix} \ = (I \mid A^{-1})$$

Thus 
$$\begin{pmatrix} 3 & -2 & 0 \\ 1 & 0 & 1 \\ -2 & 3 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{3}{5} & 0 & \frac{2}{5} \\ \frac{2}{5} & 0 & \frac{3}{5} \\ -\frac{3}{5} & 1 & -\frac{2}{5} \end{pmatrix}.$$

That is,

$$\begin{pmatrix} 3 & -2 & 0 \\ 1 & 0 & 1 \\ -2 & 3 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{5} & 0 & \frac{2}{5} \\ \frac{2}{5} & 0 & \frac{3}{5} \\ -\frac{3}{5} & 1 & -\frac{2}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} \frac{3}{5} & 0 & \frac{2}{5} \\ \frac{2}{5} & 0 & \frac{3}{5} \\ -\frac{3}{5} & 1 & -\frac{2}{5} \end{pmatrix} \begin{pmatrix} 3 & -2 & 0 \\ 1 & 0 & 1 \\ -2 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

### Why does it work?

Converting the matrix 
$$(A \mid I) = \begin{pmatrix} 3 & -2 & 0 \mid 1 & 0 & 0 \\ 1 & 0 & 1 \mid 0 & 1 & 0 \\ -2 & 3 & 0 \mid 0 & 0 & 1 \end{pmatrix}$$

into reduced row echelon form is equivalent to converting three matrices

$$\begin{pmatrix} 3 & -2 & 0 & | & 1 \\ 1 & 0 & 1 & | & 0 \\ -2 & 3 & 0 & | & 0 \end{pmatrix}, \begin{pmatrix} 3 & -2 & 0 & | & 0 \\ 1 & 0 & 1 & | & 1 \\ -2 & 3 & 0 & | & 0 \end{pmatrix}, \begin{pmatrix} 3 & -2 & 0 & | & 0 \\ 1 & 0 & 1 & | & 0 \\ -2 & 3 & 0 & | & 1 \end{pmatrix}$$

The latter are augmented matrices of certain systems of linear equations. In the matrix form,  $A\mathbf{x} = \mathbf{e}_1$ ,  $A\mathbf{x} = \mathbf{e}_2$ , and  $A\mathbf{x} = \mathbf{e}_3$ , where  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$  are consecutive columns of *I*. Suppose column vectors  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  are solutions of these systems and let  $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ . Then

$$AB = A(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = (A\mathbf{v}_1, A\mathbf{v}_2, A\mathbf{v}_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = I.$$

Row reduction and matrix algebra

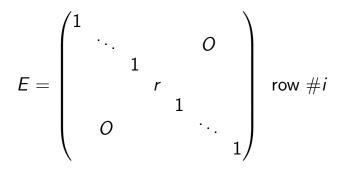
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ 2b_1 & 2b_2 & 2b_3 \\ c_1 & c_2 & c_3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 + 3a_1 & b_2 + 3a_2 & b_3 + 3a_3 \\ c_1 & c_2 & c_3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 + 3a_1 & b_2 + 3a_2 & b_3 + 3a_3 \\ c_1 & c_2 & c_3 \end{pmatrix},$$

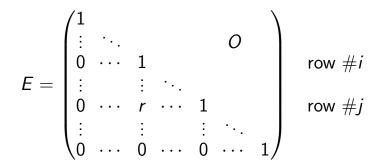
**Theorem** Any elementary row operation can be simulated as left multiplication by a certain matrix.

#### **Elementary matrices**



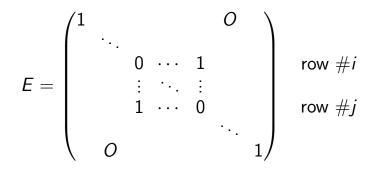
To obtain the matrix EA from A, multiply the *i*th row by r. To obtain the matrix AE from A, multiply the *i*th column by r.

#### **Elementary matrices**



To obtain the matrix EA from A, add r times the *i*th row to the *j*th row. To obtain the matrix AE from A, add r times the *j*th column to the *i*th column.

#### **Elementary matrices**



To obtain the matrix EA from A, interchange the *i*th row with the *j*th row. To obtain AE from A, interchange the *i*th column with the *j*th column.

# Why does it work? (continued)

Assume that a square matrix A can be converted to the identity matrix by a sequence of elementary row operations. Then  $E_k E_{k-1} \dots E_2 E_1 A = I$ , where  $E_1, E_2, \dots, E_k$  are elementary matrices simulating those operations.

Applying the same sequence of operations to the identity matrix, we obtain the matrix

$$B = E_k E_{k-1} \dots E_2 E_1 I = E_k E_{k-1} \dots E_2 E_1.$$

Thus BA = I. Besides, *B* is invertible since elementary matrices are invertible. It follows that  $A = B^{-1}$ , then  $B = A^{-1}$ .