MATH 304

Linear Algebra

Lecture 7: 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 $I_m A = AI_n = A$.

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

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

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 $(AB)^{-1} = B^{-1}A^{-1}$.

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

- Similarly, $(A_1A_2 \ldots A_k)^{-1} = A_k^{-1} \ldots A_2^{-1}A_1^{-1}$.

Inverting diagonal matrices

Theorem A diagonal matrix $D = \mathrm{diag}(d_1, \ldots, d_n)$ 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 $D^{-1} = \mathrm{diag}(d_1^{-1}, \ldots, d_n^{-1})$.

$$\begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_n \end{pmatrix}^{-1} = \begin{pmatrix} d_1^{-1} & 0 & \ldots & 0 \\ 0 & d_2^{-1} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_n^{-1} \end{pmatrix}$$

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

System of n linear equations in n variables:

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

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 a square matrix A, the following 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) the row echelon form of A has no zero rows;
(iv) 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}.

Theorem 3 For any $n \times n$ matrices A and B,

$$BA = I \iff AB = I.$$

Sketch of the proof: Assume $BA = I$. Then $A\mathbf{x} = \mathbf{0} \implies B(A\mathbf{x}) = B\mathbf{0} \implies (BA)\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$. By Theorem 1, A is invertible.

Row echelon form of a square matrix:

$$\begin{pmatrix} \boxed{} & * & * & * & * & * & * \\ & \boxed{} & * & * & * & * & * \\ & & \boxed{} & * & * & * & * \\ & & & \boxed{} & * & * & * \\ & & & & \boxed{} & * & * \\ & & & & & \boxed{} & * \\ & & & & & & \boxed{} \end{pmatrix}$$

invertible case

$$\begin{pmatrix} \boxed{} & * & * & * & * & * & * \\ & \boxed{} & * & * & * & * & * \\ & & \boxed{} & * & * & * & * \\ & & & & \boxed{} & * & * \\ & & & & & \boxed{} & * \\ & & & & & & \\ & & & & & & \end{pmatrix}$$

noninvertible case

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:

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

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:

$$\begin{pmatrix} \boxed{1} & 0 & 1 \\ 0 & \boxed{1} & 1.5 \\ 0 & 0 & \boxed{1} \end{pmatrix}$$

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 3×6 matrix $(A\,|\,I)$, and apply elementary row operations to this new matrix.

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

$$(A\,|\,I) = \left(\begin{array}{rrr|rrr} 3 & -2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 3 & -2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{array}\right)$$

Interchange the 1st row with the 2nd row:

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 0 & 1 & 0 \\ 3 & -2 & 0 & 1 & 0 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{array}\right)$$

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

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -2 & -3 & 1 & -3 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{array}\right)$$

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

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -2 & -3 & 1 & -3 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{array}\right)$$

Multiply the 2nd row by -0.5:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1.5 & -0.5 & 1.5 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{array}\right)$$

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

$$\left(\begin{array}{ccr|crc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1.5 & -0.5 & 1.5 & 0 \\ 0 & 0 & -2.5 & 1.5 & -2.5 & 1 \end{array}\right)$$

Multiply the 3rd row by -0.4:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1.5 & -0.5 & 1.5 & 0 \\ 0 & 0 & 1 & -0.6 & 1 & -0.4 \end{array}\right)$$

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

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0.4 & 0 & 0.6 \\ 0 & 0 & 1 & -0.6 & 1 & -0.4 \end{array}\right)$$

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

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 1 & 0 & 0.4 & 0 & 0.6 \\ 0 & 0 & 1 & -0.6 & 1 & -0.4 \end{array}\right) = (I \mid A^{-1})$$

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 $(A \,|\, I) = \left(\begin{array}{ccc|ccc} 3 & -2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{array}\right)$

into reduced row echelon form is equivalent to converting three matrices

$$\left(\begin{array}{ccc|c} 3 & -2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 \end{array}\right), \left(\begin{array}{ccc|c} 3 & -2 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ -2 & 3 & 0 & 0 \end{array}\right), \left(\begin{array}{ccc|c} 3 & -2 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -2 & 3 & 0 & 1 \end{array}\right).$$

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

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