Linear Algebra & Differential Equations

Elif Uskuplu

If you see any mistake, please email me (euskuplu@usc.edu).

---

3 Week 3

3.1 Inverse of a square matrix

In classical algebra, division is considered as the inverse operation of multiplication. For example, the multiplicative inverse of 4 is $\frac{1}{4}$ and multiplying k with $\frac{1}{4}$ is the same as dividing k by 4. The same duality occurs between addition and subtraction. The additive inverse of 4 is -4, and adding k and -4 means subtraction $k-4$.

We use this "inversion" idea to solve equations. If we have $4x = 12$, by multiplying the equation with the *inverse of* 4, we get

$$x = 1x = \frac{1}{4}4x = \frac{1}{4}12 = 3.$$

We apply the same idea in system of linear equations. Suppose we have a system $A\mathbf{x} = \mathbf{b}$, and "somehow" we have another matrix B such that $AB = BA = I_n$ (identity matrix, namely the unit of matrix multiplication) and hence we get

$$\mathbf{x} = I_n\mathbf{x} = BA\mathbf{x} = B\mathbf{b}.$$

In other words, a solution to the system is given by $B\mathbf{b}$. We consider B as an *inverse of* A.

Note that in order to have $AB = BA = I_n$, both A and B must be $n \times n$ matrices. To sum up, we will define the notion of *inverse matrix* for square matrices, and we will use it to solve the equations and to do more.

Remark. If we have $AB = BA = I_n$ and $CA = AC = I_n$, then we get

$$C = CI_n = CAB = I_nB = B.$$

In other words, if two matrices behave like *inverses* of A, they must be the same. We can say there is **the** inverse of A.

Definition 3.1. *Let A be an $n \times n$ matrix. If there exists an $n \times n$ matrix A^{-1} such that*

$$AA^{-1} = A^{-1}A = I_n,$$

then we call A^{-1} the ***inverse*** *of A. If A has the inverse, we say A is invertible.*

The followings are properties about the inverse matrix.

Theorem 3.1. *If A^{-1} exists, then the $n \times n$ system $A\boldsymbol{x} = \boldsymbol{b}$ has the unique solution given by $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ for any $\boldsymbol{b}$.*

Proof. In Section 3.5. □

Theorem 3.2. *An $n \times n$ matrix A is invertible if and only if $rank(A) = n$.*

Proof. In Section 3.5. □

Proposition 3.3. *Let A and B be invertible $n \times n$ matrices. Then*

1. *A^{-1} is invertible and $(A^{-1})^{-1} = A$.*
2. *AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*
3. *A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.*

Proof. 1. Since $A^{-1}A = I_n = AA^{-1}$, the result is trivial.

2. This is obtained by the following equalities. You can show the other equality, similarly.
$$ABB^{-1}A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

3. This is obtained by the following equalities. You can show the other equality, similarly.
$$A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n.$$

□

Theorem 3.4. *Let A and B be $n \times n$ matrices. If $AB = I_n$, then both A and B are invertible and $B = A^{-1}$.*

Proof. Page 176 of the textbook. □

Corollary 3.4.1. *Let A and B be $n \times n$ matrices. If AB is invertible, then both A and B are invertible.*

Proof. Exercise. □

3.2 Gauss-Jordan Technique to find the inverse

Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n$ be the column vectors of the identity matrix I_n. Suppose A is an $n \times n$ matrix with $rank(A) = n$, so the inverse of A exists. We want to calculate A^{-1}. If $\mathbf{x}_i$ is the i^{th} column of A^{-1}, then we get

$$A\mathbf{x}_i = \mathbf{e}_i.$$

It means that we need to solve this equation to get the column $\mathbf{x}_i$ of the inverse A^{-1}. After applying Gauss-Jordan elimination on $[A \mid \mathbf{e}_i]$, since $rank(A) = n$, we get the identity matrix on the left side, and the vector $\mathbf{x}_i$ on the right side. In other words, the reduced form of the augmented matrix is $[I_n \mid \mathbf{x}_i]$. This idea can be applied for any column of A^{-1}. Therefore, we can do all steps at once. Namely, we can start with the extended augmented matrix

$$\begin{bmatrix} A & \mid & I_n \end{bmatrix}.$$

After reducing this into the reduced row-echelon form, we will have

$$\begin{bmatrix} I_n & \mid & A^{-1} \end{bmatrix}.$$

This method to finding A^{-1} is called the **Gauss-Jordan technique**.

Example. $A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 11 \\ 4 & -3 & 10 \end{bmatrix}$ We will find the inverse A^{-1} using Gauss-Jordan method. We will reduce

$$\begin{bmatrix} 1 & -1 & 2 & \mid & 1 & 0 & 0 \\ 2 & 1 & 11 & \mid & 0 & 1 & 0 \\ 4 & -3 & 10 & \mid & 0 & 0 & 1 \end{bmatrix}$$

After applying $A_{12}(-2)$, $A_{13}(-4)$, $M_2(\frac{1}{3})$, $A_{23}(-1)$, $M_3(-3)$, $A_{32}(-\frac{7}{3})$, $A_{31}(-2)$, and $A_{21}(1)$, respectively, we get

$$\begin{bmatrix} 1 & 0 & 0 & \mid & -43 & -4 & 13 \\ 0 & 1 & 0 & \mid & -24 & -2 & 7 \\ 0 & 0 & 1 & \mid & 10 & 1 & -3 \end{bmatrix}.$$

Thus, the inverse is obtained $A^{-1} = \begin{bmatrix} -43 & -4 & 13 \\ -24 & -2 & 7 \\ 10 & 1 & -3 \end{bmatrix}$.

Exercise. Compute AA^{-1} and $A^{-1}A$ to verify that both are I_n.

3.3 Elementary matrices

Definition 3.2. *Any matrix obtained by performing a single elementary row operation on the identity matrix is called an **elementary matrix**. These are **(**P_{ij}**) Permutation Matrix:** A matrix that results from swapping two rows of* I_n*, **(**$M_i(k)$**) Scaling Matrix:** A matrix obtained by multiplying* i^{th} *row of* I_n *by a nonzero scalar, and **(**$A_ij(k)$**) Row Combination Matrix:** A matrix that results from adding a* k *multiple of* i^{th} *row to* j^{th} *row in* I_n*.*

Example. Here is the complete list of elementary matrices of size 3×3:

$$P_{12} = P_{21} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_{13} = P_{31} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad P_{23} = P_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$M_1(k) = \begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad M_2(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad M_3(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix}$$

$$A_{12}(k) = \begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_{13}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}, \quad A_{23}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$$

$$A_{21}(k) = \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_{31}(k) = \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_{32}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}$$

Remark. Premultiplying an $n \times p$ matrix A by an $n \times n$ elementary matrix E has the effect of performing the corresponding elementary row operation on A.

Example. Applying $A_{12}(5)$ on $\begin{bmatrix} 1 & 3 \\ 4 & 7 \end{bmatrix}$ yields $\begin{bmatrix} 1 & 3 \\ 9 & 22 \end{bmatrix}$. The elementary matrix $A_{12}(5)$ is a 2×2 matrix $\begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix}$. This is also obtained as

$$\begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 4 & 7 \end{bmatrix} = \begin{bmatrix} 1+0 & 3+0 \\ 5+4 & 15+7 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 9 & 22 \end{bmatrix}.$$

Remark. Since any row operation is reversible, any elementary matrix is invertible.

- The inverse of permutation matrix is $(P_{ij})^{-1} = P_{ji}$.
- The inverse of scaling matrix is $(M_i(k))^{-1} = M_i(\frac{1}{k})$.
- The inverse of row combination matrix is $(A_{ij}(k))^{-1} = A_{ij}(-k)$.

If A is an invertible $n \times n$ matrix, then reducing it to reduced row echelon form yields I_n. So it means that there are elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k E_{k-1} \ldots E_1 A = I_n.$$

Therefore, we can take $A^{-1} = E_k E_{k-1} \ldots E_1$. Also, we have

$$A = (A^{-1})^{-1} = (E_k E_{k-1} \ldots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \ldots E_{k-1}^{-1} E_k^{-1}.$$

Example. We find the inverse of $\begin{bmatrix} 1 & 3 \\ 1 & 5 \end{bmatrix}$. In order to reduce it to the identity matrix, we should apply $A_{12}(-1)$, $M_2(\frac{1}{2})$, and $A_{21}(-3)$, respectively. But it means we have

$$A_{21}(-3)M_2(\frac{1}{2})A_{12}(-1)A = I_2,$$

namely,

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

So the inverse is

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

Theorem 3.5. *An $n \times n$ matrix A is invertible iff A is a product of elementary matrices.*

Proof. Exercise. □

3.4 LU Decomposition of an invertible matrix

!!!You can skip this section if you are not interested in.

Although it is very essential concept, this section is a digression from the course. We will cover only the idea and explain its advantages.

Let A be an invertible $n \times n$ matrices. The **LU** decomposition of A means that we can write A as the product of a lower triangular and an upper triangular matrices. In other words, we have

$$A = LU$$

where L is an $n \times n$ lower triangular matrix and U is an $n \times n$ upper triangular matrix. We omit the algorithm that gives the decomposition, we focus on its usage instead.

Consider the $n \times n$ system of linear equation $A\mathbf{x} = \mathbf{b}$, where $A = LU$. If we write the system as $LU\mathbf{x} = \mathbf{b}$ and let $U\mathbf{x} = \mathbf{y}$, then solving $A\mathbf{x} = \mathbf{b}$ is equivalent to solving the pair of equations

$$L\mathbf{y} = \mathbf{b},$$

$$U\mathbf{x} = \mathbf{y}.$$

Due to the triangular form of each of coefficient matrices L and U, these systems can be solved easily by substitution.

Example. Consider the system

$$\begin{bmatrix} 6 & 18 & 3 \\ 2 & 12 & 1 \\ 4 & 15 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 19 \\ 0 \end{bmatrix}.$$

The LU decomposition of the coefficient matrix is given by

$$\begin{bmatrix} 6 & 18 & 3 \\ 2 & 12 & 1 \\ 4 & 15 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 6 & 0 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

First, let

$$\begin{bmatrix} 2 & 6 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}. \quad (*)$$

Via substitution, we can solve

$$\begin{bmatrix} 3 & 0 & 0 \\ 1 & 6 & 0 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 19 \\ 0 \end{bmatrix}.$$

Indeed, we have $y_1 = 1$, $y_2 = 3$, and $y_3 = -11$. Then, we can solve ($*$) and have $x_3 = -11$, $x_2 = 3$, and $x_1 = -6$.

It might seem that there is no advantage to using LU factorization for solving the system over Gaussian elimination. But this is true only when we try to solve a single system. If we have a set of constants $\{\mathbf{b}_i\}$ and we expect to solve $A\mathbf{x} = \mathbf{b}_i$ for each i, instead of applying a separate Gaussian elimination for each i, we can use the same LU decomposition for each $\mathbf{b}_i$. This reduces the memory storage we need in a computation. That's why many computing programs like MATLAB, NumPy (Python), Eigen (C++), R, etc. use LU decomposition as a primary algorithm to solve linear equation systems.

3.5 The Invertible Matrix Theorem

In this section, we collect all conditions for being an invertible matrix in a single theorem. To decide whether A is invertible, we can use any of the statements 2-6.

Theorem 3.6 (Invertible Matrix Theorem)**.** *Let A be an $n \times n$ matrix. The following conditions on A are equivalent:*

1. *A is invertible.*
2. *The equation $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for every $\boldsymbol{b}$ in $\mathbb{R}^n$.*
3. *The equation $A\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution $\boldsymbol{x} = \boldsymbol{0}$.*
4. *$rank(A) = n$.*
5. *A can be expressed as a product of elementary matrices.*
6. *A is row-equivalent to I_n.*

Proof. We will show each implication respectively.

$(1 \Rightarrow 2)$ For every $\mathbf{b}$, if $\mathbf{x}$ and $\mathbf{y}$ are solutions for $A\mathbf{x} = \mathbf{b}$, then we have $A\mathbf{x} = A\mathbf{y}$. Since A is invertible, we conclude

$$\mathbf{x} = I_n\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}A\mathbf{y} = I_n\mathbf{y} = \mathbf{y}.$$

Therefore, the system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

$(2 \Rightarrow 3)$ Taking $\mathbf{b} = \mathbf{0}$ we have $A\mathbf{x} = \mathbf{0}$ has a unique solution by assumption. Since the trivial solution is a solution for such system, the unique solution is the trivial solution.

$(3 \Rightarrow 4)$ Assume $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. If $rank(A) \neq n$, since $rank(A) = rank(A^{\#})$, we would have $rank(A^{\#}) < n$. It means that we would have free parameters, and hence nontrivial solutions. Since this contradicts with our assumption, we have $rank(A) = n$.

$(4 \Rightarrow 5)$ Suppose $rank(A) = n$. It means that reduced row-echelon form of A is I_n. It means that there are elementary matrices $E_1, E_2, \ldots, E_k$ such that $E_k E_{k-1} \ldots E_1 A = I_n$. Therefore, we have

$$A = (E_1^{-1}E_2^{-1} \ldots E_{k-1}^{-1}E_k^{-1})I_n = E_1^{-1}E_2^{-1} \ldots E_{k-1}^{-1}E_k^{-1}.$$

Since the inverses of elementary matrices are again elementary matrices, we get A can be expressed as a product of elementary matrices.

$(5 \Rightarrow 6)$ If A can be expressed as a product of elementary matrices $E_1, E_2, \ldots, E_k$, we have

$$A = E_1E_2 \ldots E_k = E_1E_2 \ldots E_kI_n. \quad (*)$$

However, this means that we apply corresponding elementary row operations on I_n and obtain A, which means that I_n is row equivalent to A.

$(6 \Rightarrow 1)$ If A is row equivalent to I_n, there are elementary row operations on I_n to obtain A. In other words, the equality $(*)$ holds. Since elementary matrices and I_n are invertible, and the product of invertible matrices is invertible, we conclude that A is invertible.

□

Examples using IMT.

1. Use $1 \Leftrightarrow 3$ to show that if A and B are invertible, then AB is invertible.

   Proof. In order to show AB is invertible, we will use (3) in IMT. So consider the system
   $$(AB)\mathbf{x} = \mathbf{0}. \qquad (*)$$
   Then $B\mathbf{x}$ becomes a solution for $A\mathbf{y} = \mathbf{0}$. Since A is invertible, by IMT, the last system has only trivial solution. Therefore, we get $B\mathbf{x} = \mathbf{0}$. Since B is also invertible, by the same reason, we get $\mathbf{x} = \mathbf{0}$. Since the system $(*)$ has only trivial solution, by IMT, AB is invertible. □

2. Is the statement below true or false?

   If A is a 4×4 matrix with $rank(A) = 4$, then A is row-equivalent to I_4.

   Answer. Yes, this is true by $4 \Leftrightarrow 6$ in IMT.

3. Is the statement below true or false?

   If A is a 3×3 matrix with $rank(A) = 2$, then the linear system $A\boldsymbol{x} = \boldsymbol{b}$ must have infinitely many solutions.

   Answer. False. IMT ($2 \Leftrightarrow 4$) implies there is no unique solution. But this does not mean that there are infinitely many solutions. There might be no solution. For example, consider

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}.$$

   Here, $rank(A) = 2$, but there is no solution since $rank(A^{\#}) = 3 \neq rank(A)$.

4. Is the statement below true or false?

   If the linear system $A\boldsymbol{x} = \boldsymbol{0}$ has a nontrivial solution, then A can be expressed as a product of elementary matrices.

   Answer. False because if A can be expressed as a product of elementary matrices, then by ($3 \Leftrightarrow 5$) in IMT, $A\mathbf{x} = \mathbf{0}$ would have only trivial solution.