Math 425 Applied & Comput. Lin. Algebra Fall 2024 Lecture

Matrix Inverses

We begin this lecture with a definition.

Definition 1. Let A be a square matrix of size $n \times n$. An $n \times n$ matrix X is called the inverse of A if

$$XA = AX = I_n.$$

The inverse of A is denoted by A^{-1}.

Example 1. Let $A = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$. One can check that

$$\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So we see that $A^{-1} = \begin{pmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$.

Now let's collect a few useful facts about matrix inverses.

Theorem 1. Let A be an invertible matrix. Then,

- the inverse of A is unique, i.e. there is one and only one inverse of A,
- the inverse of A is also invertible and $(A^{-1})^{-1} = A$,
- if B is also an invertible matrix of the same size as A, then their product AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$. Moreover, if $A_1, A_2, \ldots, A_k$ are invertible matrices of the same size then $(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1} A_1^{-1}$.

Proof. We prove the three statements in order.

- Such a uniqueness is usually proved like this. We will suppose that both X and Y are inverses of A, and then prove that $X = Y$. Hence, $XA = AX = I$ and $YA = AY = I$. Now $X = XI = X(AY) = (XA)Y = IY = Y$.
- This statement is not tricky as it looks. Since A^{-1} is the inverse of A we have $A^{-1}A = AA^{-1} = I$. But this says, by definition of inverses, that the inverse of A^{-1} is A.
- We just check that $(B^{-1}A^{-1})(AB) = I$ and $(AB)(B^{-1}A^{-1}) = I$. The general statement is proved similarly.

$\square$

Gauss-Jordan Elimination. Our discussion above does not show a method of finding the inverse of a matrix if it exists with the exception of 2×2 matrices and elementary matrices of Type 1 and Type 2. In this subsection we will go over a method known as *Gauss-Jordan elimination* of producing the inverse of an invertible $n \times n$ matrix A.

The idea is quite simple: if A is invertible then there is an $n \times n$ matrix X, the inverse of A and yet unknown, such that $AX = I_n$. So let $X = (\mathbf{x}_1\, \mathbf{x}_2\, \ldots \mathbf{x}_n)$ where $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n$ are the columns of X, and let

$I_n = (\mathbf{e}_1\,\mathbf{e}_2\,\ldots\mathbf{e}_n)$ where $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$ are the columns of the identity matrix I_n. Whereas the columns of X are not known we know the columns of I_n:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad \ldots \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

In order to find jth column of X we can solve the system $A\mathbf{x}_j = \mathbf{e}_j$ using Gaussian elimination. So we can set up n augmented matrices $(A|\mathbf{x}_j)$ and proceed. However, since we will do *exactly* the same row operations in each of the n Gaussian eliminations we will apply Gaussian elimination to $(A|I_n)$.

An example will help. Say we want to find the inverse of $A = \begin{pmatrix} 0 & 1 & -2 \\ 5 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix}$. We set up the augmented matrix $(A|I_3) = \left(\begin{array}{ccc|ccc} 0 & 1 & -2 & 1 & 0 & 0 \\ 5 & 3 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$ and do Gaussian elimination

$$\left(\begin{array}{ccc|ccc} 0 & 1 & -2 & 1 & 0 & 0 \\ 5 & 3 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 5 & 3 & -1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{ccc|ccc} 5 & 3 & -1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & \frac{2}{5} & \frac{6}{5} & 0 & -\frac{1}{5} & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 5 & 3 & -1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 2 & -\frac{2}{5} & -\frac{1}{5} & 1 \end{array}\right).$$

At this moment A has been reduced to an upper triangular matrix U with nonzero diagonals and we have the augmented matrix $(U\,|\,C)$ and we may solve $U\mathbf{x}_j = \mathbf{c}_j$ by back substitution (n of them!). Luckily, there is a better way to go. Let's assume for a moment that by making further row operations we were able to reduce $(U\,|\,C)$ to $(I\,|\,D)$. Once we arrive to this augmented matrix we solve $I\mathbf{x}_j = \mathbf{d}_j$ by back substitution. But wait! We do not need to do back substitution at all. We can just conclude $\mathbf{x}_j = \mathbf{d}_j$, i.e., we can read off the jth column of the matrix X. In other words when we arrived to $(I\,|\,D)$ we really arrived to $(I\,|\,X)$. Voila! We computed the inverse $X = A^{-1}$.

We explain how this procedure works on our example. Our first job is to make all the diagonal entries of U equal to one (if they are not yet so). For instance, we multiply the first row by $\frac{1}{5}$ to make the $(1,1)$ entry equal to 1. Note that this amounts to scaling an equation of a linear system and this does not change the solutions.

$$\left(\begin{array}{ccc|ccc} 5 & 3 & -1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 2 & -\frac{2}{5} & -\frac{1}{5} & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & \frac{3}{5} & -\frac{1}{5} & 0 & \frac{1}{5} & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 2 & -\frac{2}{5} & -\frac{1}{5} & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & \frac{3}{5} & -\frac{1}{5} & 0 & \frac{1}{5} & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{array}\right).$$

Before we go on it is good to stop and record this last type of elementary row operation. We call this Elementary Row Operation Type 3:

Multiply a row by a nonzero scalar c

Now we can continue Gauss-Jordan elimination. After the above scalings we have arrived at $(V\,|\,\hat{C})$ where V is upper triangular whose diagonal entries are all ones. Now we can do further elementary row operations of type to turn all nonzero entries of V above its diagonal equal to zero, arriving at the identity matrix. In our example, here is how it goes:

$$\left(\begin{array}{ccc|ccc} 1 & \frac{3}{5} & -\frac{1}{5} & 0 & \frac{1}{5} & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & -\frac{3}{5} & \frac{1}{5} & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & -\frac{3}{5} & \frac{1}{5} & 0 \\ 0 & 1 & 0 & \frac{3}{5} & -\frac{1}{5} & 1 \\ 0 & 0 & 1 & -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{5} & \frac{3}{10} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{3}{5} & -\frac{1}{5} & 1 \\ 0 & 0 & 1 & -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{array}\right).$$

Let's check:

$$AX = \begin{pmatrix} 0 & 1 & -2 \\ 5 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -\frac{2}{5} & \frac{3}{10} & -\frac{1}{2} \\ \frac{3}{5} & -\frac{1}{5} & 1 \\ -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} -\frac{2}{5} & \frac{3}{10} & -\frac{1}{2} \\ \frac{3}{5} & -\frac{1}{5} & 1 \\ -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & -2 \\ 5 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

At the end we arrive to an important result that we are ready to prove. This result says for a $n \times n$ A matrix nonsingular means the same as invertible.

Theorem 2. A $n \times n$ matrix is nonsingular if and only if it is invertible.

Proof. The proof uses Gauss-Jordan elimination. If A is nonsingular by the first part of the Gauss-Jordan elimination we can reduce $(A \,|\, I)$ to $(U \,|\, C)$ where U is upper triangular with nonzero diagonals (this is what it means to be nonsingular after all). Now we can continue Gauss-Jordan elimination to arrive $(I \,|\, X)$ where $X = A^{-1}$. This shows that A is invertible. Conversely, suppose that A is invertible. Then applying Gauss-Jordan elimination to $(A \,|\, I)$ *must* produce $(I \,|\, X)$. So we start Gauss-Jordan elimination but allow ourselves only type 1 and type 2 row operations. This must reduce the augmented matrix to $(U \,|\, C)$ eventually where U is upper triangular with nonzero entries since otherwise Gauss-Jordan could not have worked. But this means A is nonsingular. $\square$

Theorem 3. Let $A\mathbf{x} = \mathbf{b}$ be a system of linear equations where A is a $n \times n$ nonsingular matrix. Then the unique solution to this system is obtaineb by $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof. Since A is nonsingular, by the above theorem, it is also invertible. Therefore $A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b}$. The associativity of matrix multiplication gives us $(A^{-1}A)\mathbf{x} = A^{-1}\mathbf{b}$. Since $A^{-1}A = I$ we get the result. $\square$

Example 2. If we wish to solve

$$\begin{pmatrix} 0 & 1 & -2 \\ 5 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 9 \\ -1 \end{pmatrix}$$

we can use the inverse of the coefficient matrix we have computed since we are lucky enought that this matrix is invertible:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\frac{2}{5} & \frac{3}{10} & -\frac{1}{2} \\ \frac{3}{5} & -\frac{1}{5} & 1 \\ -\frac{1}{5} & -\frac{1}{10} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} -2 \\ 9 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \\ -1 \end{pmatrix}.$$

However, nobody in his right mind will solve a system first finding the inverse A of its coefficient matrix A^{-1} and *then* computing $A^{-1}\mathbf{b}$. Gaussian elimination followed by back substitution is cheaper and numerically more stable than than Gauss-Jordan elimination so we almost never compute the inverse of a matrix to solve a system.