MA1508E: LINEAR ALGEBRA FOR ENGINEERING

Lecture 3 Notes

References

1. Elementary Linear Algebra: Application Version, Section 1.4-1.6
2. Linear Algebra with Application, Section 2.4-2.5

2.5 Block Multiplication

Let $\mathbf{A}$ be an $m \times n$ matrix,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

The <u>rows</u> of $\mathbf{A}$ are the $1 \times n$ submatrices of $\mathbf{A}$,

$$\mathbf{r}_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix}, \mathbf{r}_2 = \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{pmatrix}, ..., \mathbf{r}_m = \begin{pmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

and the <u>columns</u> of $\mathbf{A}$ are the $m \times 1$ submatrices of $\mathbf{A}$

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

In general, a $p \times q$ <u>submatrix</u> of an $m \times n$ matrix $\mathbf{A}$, $p \le m$, $q \le n$, is formed by taking a $p \times q$ block of the entries of the matrix $\mathbf{A}$.

Example. $\begin{pmatrix} 4 & 6 & 1 \\ 2 & 2 & 1 \end{pmatrix}$ is a 2×3 submatrix of $\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 & 7 & 1 \\ 2 & 4 & 6 & 1 & 2 \\ 1 & 2 & 2 & 1 & 3 \end{pmatrix}$, taken from row 2 and 3, and columns 2 to 4 of $\mathbf{A}$.

Observe that matrix multiplication respects submatrices, in the sense that if we take $k \times p$ submatrices of $\mathbf{A}$ and multiply to $p \times l$ submatrices of $\mathbf{B}$, we obtain $k \times l$ submatrices of $\mathbf{AB}$. We call this <u>block multiplication</u>.

Example. Let $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 3 & 2 & 1 \\ 4 & 4 & 2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 3 & 5 & 7 & 1 \\ 2 & 4 & 6 & 1 & 2 \\ 1 & 2 & 2 & 1 & 3 \end{pmatrix}$. Then multiplying the submatrix of $\mathbf{A}$, $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix}$ consisting of the first 2 rows of $\mathbf{A}$ to the submatrix of $\mathbf{B}$,

$\begin{pmatrix}1 & 3\\ 2 & 4\\ 1 & 2\end{pmatrix}$ consisting of the first 2 columns of $\mathbf{B}$, we get $\begin{pmatrix}2 & 5\\ 7 & 15\end{pmatrix}$, which is a 2×2 submatrix of $\mathbf{AB}$ consisting of the first 2 rows and first 2 columns of $\mathbf{AB}$.

$$\begin{pmatrix}1 & 0 & 1\\ 1 & 2 & 2\\ 3 & 2 & 1\\ 4 & 4 & 2\end{pmatrix}\begin{pmatrix}1 & 3 & 5 & 7 & 1\\ 2 & 4 & 6 & 1 & 2\\ 1 & 2 & 2 & 1 & 3\end{pmatrix} = \begin{pmatrix}2 & 5 & 7 & 8 & 4\\ 7 & 15 & 21 & 11 & 11\\ 8 & 9 & 29 & 24 & 10\\ 14 & 32 & 48 & 34 & 18\end{pmatrix}.$$

In particular, we have the following cases.

1. If $\mathbf{B} = \begin{pmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n\end{pmatrix}$, then
$$\mathbf{AB} = \begin{pmatrix}\mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_n\end{pmatrix},$$
that is, the columns of $\mathbf{AB}$ is just $\mathbf{A}$ pre-multiplying to the columns of $\mathbf{B}$.

2. If $\mathbf{A} = \begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\\ \vdots\\ \mathbf{a}_m\end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n\end{pmatrix}$, then
$$\mathbf{AB} = \begin{pmatrix}\mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_n\\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_n\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{a}_m\mathbf{b}_1 & \mathbf{a}_m\mathbf{b}_2 & \cdots & \mathbf{a}_m\mathbf{b}_n\end{pmatrix}.$$
Indeed, the (i,j)-entry of $\mathbf{AB}$ is
$$\mathbf{a}_i\mathbf{b}_j = \begin{pmatrix}a_{i1} & a_{i2} & \cdots & a_{ip}\end{pmatrix}\begin{pmatrix}b_{ij}\\ b_{2j}\\ \vdots\\ b_{pj}\end{pmatrix} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

Exercise: If $\mathbf{A}_i$ is a $m_i \times p$ matrix, for $i = 1, 2$, and $\mathbf{B}_i$ is a $p \times n_i$ matrix for $i = 1, 2$, show that the following block multiplication holds,

$$\begin{pmatrix}\mathbf{A}_1\\ \mathbf{A}_2\end{pmatrix}\begin{pmatrix}\mathbf{B}_1 & \mathbf{B}_2\end{pmatrix} = \begin{pmatrix}\mathbf{A}_1\mathbf{B}_1 & \mathbf{A}_1\mathbf{B}_2\\ \mathbf{A}_2\mathbf{B}_1 & \mathbf{A}_2\mathbf{B}_2\end{pmatrix},$$

where $\begin{pmatrix}\mathbf{A}_1\\ \mathbf{A}_2\end{pmatrix}$ is a $(m_1 + m_2) \times p$ matrix, where the first m_1 rows are from $\mathbf{A}_1$, and the $m_1 + 1$ to $m_1 + m_2$ rows are from $\mathbf{A}_2$, and $\begin{pmatrix}\mathbf{B}_1 & \mathbf{B}_2\end{pmatrix}$ is a $p \times (n_1 + n_2)$ matrix, where the first n_1 columns are from $\mathbf{B}_1$ and the $n_1 + 1$ to $n_1 + n_2$ columns are from $\mathbf{B}_2$.

Suppose now we have p linear systems with the same coefficient matrix $\mathbf{A} = (a_{ij})_{m\times n}$, for $k = 1, ..., p$,

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_{1k}\\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_{2k}\\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_{mk}\end{array}.$$

We can represent this as a matrix equation $\mathbf{AX} = \mathbf{B}$, where $\mathbf{X} = (x_{ij})_{n\times p}$, and $\mathbf{B} = (b_{ij})_{m\times p}$,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mp} \end{pmatrix}.$$

This can be reprenseted as a p simultaneous augmented matrix

$$\left(\begin{array}{cccc|c|c|c|c} a_{11} & a_{12} & \cdots & a_{1n} & b_{11} & b_{12} & & b_{1p} \\ a_{21} & a_{22} & \cdots & a_{2n} & b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_{m1} & b_{m2} & & b_{mp} \end{array}\right).$$

Then we may proceed to perform row operations to the augmented matrix above, and solve for all p linear systems simultaneously.

Example. 1. Solve the following 2 linear systems

$$\begin{array}{rcrcrcr} x & + & 2y & - & 3z & = & 1 \\ 2x & + & 6y & - & 11z & = & 1 \\ x & - & 2y & + & 7z & = & 1 \end{array} \quad \text{and} \quad \begin{array}{rcrcrcr} x & + & 2y & - & 3z & = & 1 \\ 2x & + & 6y & - & 11z & = & 2 \\ x & - & 2y & + & 7z & = & 1 \end{array}$$

can be represented as a matrix equation

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 6 & -11 \\ 1 & -2 & 7 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \\ z_1 & z_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \end{pmatrix},$$

and the augmented matrix is

$$\left(\begin{array}{ccc|c|c} 1 & 2 & -3 & 1 & 1 \\ 2 & 6 & -11 & 1 & 2 \\ 1 & -2 & 7 & 1 & 1 \end{array}\right).$$

The reduced row-echelon form is

$$\left(\begin{array}{ccc|c|c} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -5/2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

So the first linear system is inconsistent, and the second has a general solution $x = 1 - 2s, y = (5/2)s, z = s$.

2. Solve the following 3 linear systems

$$\begin{array}{rcrcrcr} 3x & + & 2y & - & z & = & a \\ 5x & - & y & + & 3z & = & b \\ 2x & + & y & - & z & = & c \end{array}, \text{ for } \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

The matrix equation is

$$\begin{pmatrix} 3 & 2 & -1 \\ 5 & -1 & 3 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{pmatrix},$$

and the augmented matrix is

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

The reduced row-echelon form is

$$\left(\begin{array}{ccc|c|c|c} 1 & 0 & 0 & 5/3 & 2/9 & -1/9 \\ 0 & 1 & 0 & -11/3 & 7/9 & 10/9 \\ 0 & 0 & 1 & -10/3 & 2/9 & 8/9 \end{array}\right).$$

Hence, all 3 linear systems has unique solutions,

$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 5 \\ -11 \\ 10 \end{pmatrix}, \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 2 \\ 7 \\ 2 \end{pmatrix}, \begin{pmatrix} x_{13} \\ x_{13} \\ x_{33} \end{pmatrix} = \frac{1}{9} \begin{pmatrix} -1 \\ 10 \\ 8 \end{pmatrix}.$$

2.6 Inverse of a Matrix

Recall that for a nonzero real number $c \in \mathbb{R}$, the multiplicative inverse (or just inverse) of c, denoted as $\frac{1}{c}$, is defined to be the number such that when multiplied to c gives 1, $\frac{1}{c} \times c = 1$ (this is why 0 has no inverse, for $a \times 0 = 0$ for any real number $a \in \mathbb{R}$, and so there can be no $a \in \mathbb{R}$ such that $a \times 0 = 1$).

Hence, in order to define, if possible, the inverse of a matrix, we first need to identity the matrix that serves the same role as 1 in real numbers. We indeed have such an object. Recall that the identity matrix has the multiplicative identity property, for any $m \times n$ matrix $\mathbf{A}$,

$$\mathbf{I}_m \mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_n.$$

Remark. We may define an inverse of a $m \times n$ matrix $\mathbf{A}$ to be a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$. However, since matrix multiplication is non-commutative, we might not have $\mathbf{AB} = \mathbf{I}$. In fact, if $m \neq n$, then for $\mathbf{AB}$ and $\mathbf{BA}$ to be defined, $\mathbf{B}$ must be of size $n \times m$, and so $\mathbf{AB}$ is not even of the same size as $\mathbf{BA}$. Therefore, we shall consider them separately.

Let $\mathbf{A}$ be a $m \times n$ matrix. A <u>left inverse</u> of $\mathbf{A}$ is a matrix $n \times m$ matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}_n$. A <u>right inverse</u> of $\mathbf{A}$ is a $n \times m$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}_m$.

Example. 1. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

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

Remark. 1. Note that if $\mathbf{B}$ is a left inverse of $\mathbf{A}$, then $\mathbf{A}$ is a right inverse of $\mathbf{B}$.

2. In the definition, we need not specify the size of the left or right inverse $\mathbf{B}$, since identity matrices are square matrices, the size of $\mathbf{B}$ is determined by $\mathbf{A}$.

3. Not every matrix has a left or right inverse. For example, show that $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ cannot have left or right inverse. More generally, if $n > m$, then any $m \times n$ matrix cannot have a left inverse, and if $m > n$, then any $m \times n$ matrix cannot have a right inverse. The proof of this statement is postponed till lecture 9.

4. Left and right inverses are not unique. For example,

$$\begin{pmatrix} 1 & 0 & s \\ 0 & 1 & t \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ a & b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

for any $a, b, s, t \in \mathbb{R}$.

5. For this reason and more, we will not delve too much into inverses for non square matrices. It turns out that if we restrict our attention to square matrices, we have (almost) all the desired properties that inverses of real numbers have.

A square matrix $\mathbf{A}$ of order n is <u>invertible</u> if there exists a square matrix $\mathbf{B}$ of order n such that $\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}$. A square matrix is <u>singular</u> if it is not invertible.

Remark. 1. As before, it is not required to specify in the definition the size of $\mathbf{B}$, it is a consequence of the definition.

2. For $\mathbf{A}$ to be invertible, we need to check that there exists a $\mathbf{B}$ that is simultaneously a left and right inverse of $\mathbf{A}$, that is, we need to show both $\mathbf{BA} = \mathbf{I}_n$ and $\mathbf{AB} = \mathbf{I}_n$.

3. It turns out that to show that a square matrix $\mathbf{A}$ is invertible, suffice to show that it has a left inverse, or a right inverse. Moreoever, if $\mathbf{BA} = \mathbf{I}$, then neessarily $\mathbf{AB} = \mathbf{I}$ and vice versa. The actual theorem and proof will be given in the next lecture.

Theorem (Uniqueness of inverse)**.** *If* $\mathbf{B}$ *and* $\mathbf{C}$ *are both inverses of a square matrix* $\mathbf{A}$*, then* $\mathbf{B} = \mathbf{C}$*.*

Proof. By definition, $\mathbf{BA} = \mathbf{I}_n = \mathbf{AC}$. So

$$\mathbf{B} = \mathbf{BI} = \mathbf{B}(\mathbf{AC}) = (\mathbf{BA})\mathbf{C} = \mathbf{IC} = \mathbf{C}.$$

□

So since inverses are unique, we can denote the inverse of an invertible matrix $\mathbf{A}$ by $\mathbf{A}^{-1}$. That is, if $\mathbf{A}$ is invertible, there exists a unique matrix $\mathbf{A}^{-1}$ such that

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

Example. 1. The identity matrix $\mathbf{I}$ is invertible with $\mathbf{I}^{-1} = \mathbf{I}$.

2. $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. So $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$.

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

Theorem (Cancellation law for matrices)**.** *Let* $\mathbf{A}$ *be a invertible matrix of order* n*.*

(i) If $\mathbf{B}$ *and* $\mathbf{C}$ *are* $n \times m$ *matrices with* $\mathbf{AB} = \mathbf{AC}$*, then* $\mathbf{B} = \mathbf{C}$*.*

(ii) If $\mathbf{B}$ *and* $\mathbf{C}$ *are* $m \times n$ *matrices with* $\mathbf{BA} = \mathbf{CA}$*, then* $\mathbf{B} = \mathbf{C}$*.*

Proof. (i) Pre-multiply $\mathbf{A}^{-1}$ to both sides of $\mathbf{AB} = \mathbf{AC}$, we get $\mathbf{B} = \mathbf{IB} = \mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{AC} = \mathbf{IC} = \mathbf{C}$.

(ii) Post-multiply $\mathbf{A}^{-1}$ to both sides of $\mathbf{BA} = \mathbf{CA}$, we get we get $\mathbf{B} = \mathbf{BI} = \mathbf{BAA}^{-1} = \mathbf{CAA}^{-1} = \mathbf{CI} = \mathbf{C}$.

□

Exercise:

1. Let $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$. By writing $\mathbf{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, using the equation $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$, find the inverse of $\mathbf{A}$. What do you notice about the equations derived from $\mathbf{AB} = \mathbf{I}$ and $\mathbf{I} = \mathbf{BA}$?

2. Show that if $\mathbf{A}$ is invertible, the linear system $\mathbf{Ax} = \mathbf{b}$ is consistent.

Theorem (Properties of inverses)**.** *Let* $\mathbf{A}$ *be an invertible matrix of order* n*.*

(i) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

(ii) For any nonzero real number $a \in \mathbb{R}$*,* $(a\mathbf{A})$ *is invertible with inverse* $(a\mathbf{A})^{-1} = \frac{1}{a}\mathbf{A}^{-1}$.

(iii) $\mathbf{A}^T$ *is invertible with inverse* $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

(iv) If $\mathbf{B}$ *is an invertible matrix of order* n*, then* $(\mathbf{AB})$ *is invertible with inverse* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

Proof. (i) Since $\mathbf{AA}^{-1} = \mathbf{I}_n = \mathbf{A}^{-1}\mathbf{A}$, then $\mathbf{A}$ is the inverse of $\mathbf{A}^{-1}$.

(ii) We directly check that $a\mathbf{A}(\frac{1}{a}\mathbf{A}^{-1}) = \frac{a}{a}\mathbf{I}_n = \mathbf{I}_n = (\frac{1}{a}\mathbf{A}^{-1})(a\mathbf{A})$.

(iii) Since $\mathbf{I}$ is symmetric, $\mathbf{I} = \mathbf{I}^T = (\mathbf{AA}^{-1})^T = (\mathbf{A}^{-1})^T\mathbf{A}^T$, and similarly, $\mathbf{I} = \mathbf{A}^T(\mathbf{A}^{-1})^T$. Hence, $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

(iv) We directly check that $(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}_n$ and $(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}_n$

□

Remark. 1. By induction, one can prove that the product of invertible matrices is invertible, and $(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\cdots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}$ if $\mathbf{A}_i$ is an invertible matrix for $i = 1, ..., k$.

2. We define the negative power of an invertible matrix to be

$$\mathbf{A}^{-n} = (\mathbf{A}^{-1})^n$$

for any positive integer n.

Exercise: Show that $\mathbf{AB}$ is invertible if and only if both $\mathbf{A}$ and $\mathbf{B}$ are invertible. Hint: If $\mathbf{AB}$ is invertible, let $\mathbf{C}$ be the inverse. Pre and post multiply $\mathbf{AB}$ with $\mathbf{C}$.

Theorem. *An order 2 square matrix*

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible if and only if $ad - bc \neq 0$, in which case the inverse is given by the formula

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The formula is obtained by using the adjoint of $\mathbf{A}$, which is beyond the scope of this module. However, readers may verify that indeed we have $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_2 = \mathbf{A}^{-1}\mathbf{A}$.

2.7 Elementary Matrices

A square matrix of order n $\mathbf{E}$ is called an elementary matrix if it can be obtained from the identity matrix $\mathbf{I}_n$ by performing a single elementary row operation

$$\mathbf{I}_n \xrightarrow{r} \mathbf{E},$$

where r is an elementary row operation. The elementary row operations is said to be the row operation corresponding to the elementary matrix.

Theorem (Elementary matrices and elementary row operations)**.** *Let $\mathbf{A}$ be an $n \times m$ matrix and let $\mathbf{E}$ be the elementary matrix*

(i)

$$\begin{array}{cc} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & 0 & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & & a & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \\ & \begin{matrix} & & i & & j & & \end{matrix} \end{array} \text{ if } i < j, \text{ or } \begin{array}{cc} \begin{matrix} \\ \\ j \\ \\ i \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & 0 & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & a & & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \\ & \begin{matrix} & & i & & j & & \end{matrix} \end{array} \text{ if } i > j,$$

(ii)

$$\begin{array}{cc} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & 0 & & & & & 0 \\ & \ddots & & & & & \\ 0 & & & & 1 & & 0 \\ & & & \ddots & & & \\ 0 & & 1 & & & & 0 \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \\ & \begin{matrix} & & i & & j & & \end{matrix} \end{array},$$

(iii)

$$\begin{array}{cc} i & \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} \\ & \begin{array}{ccccc} & & i & & \end{array} \end{array}, \quad c \neq 0.$$

Then the product $\mathbf{EA}$ *is obtained from* $\mathbf{A}$ *via performing the corresponding elementary row operations*

(i) adding a *times row* j *to row* i $(R_i + aR_j)$*,*

(ii) exchanging row i *with row* j $(R_i \leftrightarrow R_j)$*,*

(iii) multiplying row i *by* c*,*

respectively. In words, it means that pre-multiplying an elementary matrix is equivalent to performing the corresponding elementary row operation.

Proof. The reader can directly verify by computing the product $\mathbf{EA}$ for each of the three cases. □

Example. 1.

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} \xrightarrow{R_2+2R_1} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 4 & 5 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix}$$

2.

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 4 & 2 \\ 1 & 2 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix}$$

3.

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} \xrightarrow{-2R_2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ -2 & -4 & -6 & 2 \\ 2 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix}$$

Corollary. *Suppose the matrix* $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *by performing row operations* $r_1, r_2, ..., r_k$*,*

$$\mathbf{A} \xrightarrow{r_1} \xrightarrow{r_2} \cdots \xrightarrow{r_k} \mathbf{B}.$$

Let $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_k$ *be the corresponding elementary matrices. Then*

$$\mathbf{B} = \mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A}.$$

Proof. We will proof by induction. Let $\mathbf{A}_1$ be the matrix obtained by performing r_1 on $\mathbf{A}$. Then by the theorem above, $\mathbf{A}_1 = \mathbf{E}_1\mathbf{A}$. Suppose now $\mathbf{A}_l$ is the matrix obtained by performing row operations $r_1, r_2, ..., r_l$, $l < k$. By the induction hypothesis, $\mathbf{A}_l = \mathbf{E}_l \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A}$. Let $\mathbf{A}_{l+1}$ be obtained from $\mathbf{A}_l$ by performing row operation r_l. Then by the theorem above,

$$\mathbf{A}_{l+1} = \mathbf{E}_{l+1}\mathbf{A}_l = \mathbf{E}_{l+1}\mathbf{E}_l \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A}.$$

The inductive step is shown, and hence, the statement is proved. □

Example.

$$\begin{pmatrix}1&1&1&0\\1&2&3&-1\\2&1&4&2\end{pmatrix}\xrightarrow{R_2+2R_1}\xrightarrow{R_2\leftrightarrow R_3}\begin{pmatrix}1&1&1&0\\2&1&4&2\\3&4&5&-1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&1&1&0\\1&2&3&-1\\2&1&4&2\end{pmatrix}$$

Question: What if we post-multiply elementary matrices to a matrix?

Lemma. *Elementary matrix* $\mathbf{E}$ *is invertible, and the inverse* $\mathbf{E}^{-1}$ *is also an elementary matrix.*

Proof. We claim that the inverse of the elementary matrix corresponding to

(i) $R_i + aR_j$,

(ii) $R_i \leftrightarrow R_j$, and

(iii) cR_i for $c \neq 0$

is the elementary matrix corresponding to

(i) $R_i - aR_j$,

(ii) $R_i \leftrightarrow R_j$, and

(iii) $\frac{1}{c}R_i$,

respectively. The verification is left to the reader. □

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

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

3. $\begin{pmatrix}1&0&0\\0&1&0\\0&0&5\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1/5\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1/5\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&5\end{pmatrix}$

2.8 Inverses and elementary matrices

Theorem. *A square matrix is invertible if and only if its reduce row-echelon form is the identity matrix.*

Equivalently, a square matrix is singular if and only if it has a REF with a zero row.

Corollary. *A square matrix is invertible if and only if it is a product of elementary matrices.*

Example. 1.

$$\mathbf{A} = \begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 2\\ 0 & 1 & 1\end{pmatrix} \xrightarrow{R_2-R_1} \begin{pmatrix}1 & 1 & 1\\ 0 & 0 & 1\\ 0 & 1 & 1\end{pmatrix} \xrightarrow{R_2\leftrightarrow R_3} \begin{pmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix} \xrightarrow{R_1-R_2} \xrightarrow{R_2-R_3} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

So

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & -1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\mathbf{A} = \mathbf{I},$$

or, $\mathbf{A}^{-1} = \begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}.$

2.

$$\mathbf{A} = \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 1 & 1 & 1\end{pmatrix} \xrightarrow{R_3-R_1} \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 1 & 0\end{pmatrix} \xrightarrow{R_3-R_2} \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Hence, $\mathbf{A}$ is singular.

Recall that two matrices $\mathbf{A}$ and $\mathbf{B}$ are row equivalent if one can be obtained from the other by performing elementary row operations. This is equivalent to $\mathbf{B}$ being obtained from $\mathbf{A}$ by pre-multiplying some elementary row operations.

Exercise: Show that $\mathbf{A}$ and $\mathbf{B}$ are row equivalent if and only if $\mathbf{A} = \mathbf{PB}$ for some invertible matrix $\mathbf{P}$.

Algorithm for finding inverses

The theorem and corollary above provides us with a way to check if a matrix is invertible, and also to find the inverse if it exists. Let $\mathbf{A}$ be a square matrix of order n.

Step 1: Form a new $n \times 2n$ matrix $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I}_n \end{array}\right)$.

Step 2: Reduce the matrix $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array}\right) \longrightarrow \left(\begin{array}{c|c} \mathbf{R} & \mathbf{B} \end{array}\right)$ to its reduced row-echelon form.

Step 3: If $\mathbf{R} \neq \mathbf{I}$, then $\mathbf{A}$ is not invertible. If $\mathbf{R} = \mathbf{I}$, $\mathbf{A}$ is invertible with inverse $\mathbf{A}^{-1} = \mathbf{B}$.

We will explain why $\mathbf{A}^{-1} = \mathbf{B}$. Since $\left(\begin{array}{c|c} \mathbf{R} & \mathbf{B} \end{array}\right)$ is the RREF of $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I}_n \end{array}\right)$, there an invertible matrix $\mathbf{P}$ such that

$$\left(\begin{array}{c|c} \mathbf{R} & \mathbf{B} \end{array}\right) = \mathbf{P}\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{PA} & \mathbf{PI} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{PA} & \mathbf{P} \end{array}\right).$$

Since $\mathbf{A}$ is invertible, $\mathbf{R} = \mathbf{I}$ and the equation above tells us that $\mathbf{B} = \mathbf{P}$ and $\mathbf{I} = \mathbf{PA} = \mathbf{BA}$, and thus $\mathbf{A}^{-1} = \mathbf{B}$.

Example. Find the inverse of $\begin{pmatrix} 2 & 7 & 1 \\ 1 & 4 & -1 \\ 1 & 3 & 0 \end{pmatrix}$.

$$\left(\begin{array}{ccc|ccc} 2 & 7 & 1 & 1 & 0 & 0 \\ 1 & 4 & -1 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 0 & 1 \\ 1 & 4 & -1 & 0 & 1 & 0 \\ 2 & 7 & 1 & 1 & 0 & 0 \end{array}\right) \xrightarrow[R_2 - R_1]{R_3 - 2R_1}$$
$$\left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 & 0 & -2 \end{array}\right) \xrightarrow{R_2 + R_3} \xrightarrow{\frac{1}{2}R_2} \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1/2 & 1/2 & -3/2 \\ 0 & 1 & 1 & 1 & 0 & -2 \end{array}\right) \xrightarrow[R_1 - 3R_2]{R_3 - R_2}$$
$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -3/2 & -3/2 & 11/2 \\ 0 & 1 & 0 & 1/2 & 1/2 & -3/2 \\ 0 & 0 & 1 & 1/2 & -1/2 & -1/2 \end{array}\right).$$

So
$$\begin{pmatrix} 2 & 7 & 1 \\ 1 & 4 & -1 \\ 1 & 3 & 0 \end{pmatrix}^{-1} = \frac{1}{2} \begin{pmatrix} -3 & -3 & 11 \\ 1 & 1 & -3 \\ 1 & -1 & -1 \end{pmatrix}.$$

Appendix for Lecture 3

Theorem. *A square matrix is invertible if and only if its reduce row-echelon form is the identity matrix.*

Proof. Let $\mathbf{R}$ be the reduced row-echelon form of a square matrix $\mathbf{A}$ of order n. Then $\mathbf{R} = \mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{PA}$, where $\mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1$, for some elementary matrices $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_k$. Equivalently,

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}\mathbf{R}.$$

Note that $\mathbf{R}$ is the identity matrix if and only if it has n pivot columns.

Suppose $\mathbf{R}$ has n pivot columns. Then $\mathbf{R} = \mathbf{I}$, and so

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1},$$

which shows that $\mathbf{A}$ is invertible since it is a product of invertible matrices.

Suppose $\mathbf{R}$ has less than n pivot columns. Then the last row of $\mathbf{R}$ must be a zero row. Write

$$\mathbf{R} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{0}_{1\times n} \end{pmatrix},$$

for some $(n-1) \times n$ matrix $\mathbf{Q}$. Then for any square matrix $\mathbf{B}$ of order n,

$$\mathbf{PAB} = \mathbf{RB} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{0}_{1\times n} \end{pmatrix} \mathbf{B} = \begin{pmatrix} \mathbf{QB} \\ \mathbf{0}_{1\times n}\mathbf{B} \end{pmatrix} = \begin{pmatrix} \mathbf{QB} \\ \mathbf{0}_{1\times n} \end{pmatrix},$$

which can never be equal to the identity since it has a zero row. So $\mathbf{PA}$ cannot be invertible, and thus $\mathbf{A}$ cannot be invertible. $\square$

Corollary. *A square matrix is invertible if and only if it is a product of elementary matrices.*

Proof. Using the notations from the previous theorem, we have $\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}\mathbf{R}$, and by a lemma above, each $\mathbf{E}_i^{-1}$ is an elementary matrix, for $i = 1, ..., k$. By the theorem above, if $\mathbf{A}$ is invertible, $\mathbf{R} = \mathbf{I}$, and so $\mathbf{A}$ is a product of elementary matrices.

Conversely, if $\mathbf{A}$ is a product of elementary matrices, then since elementary matrices are invertible, $\mathbf{A}$ is invertible. $\square$