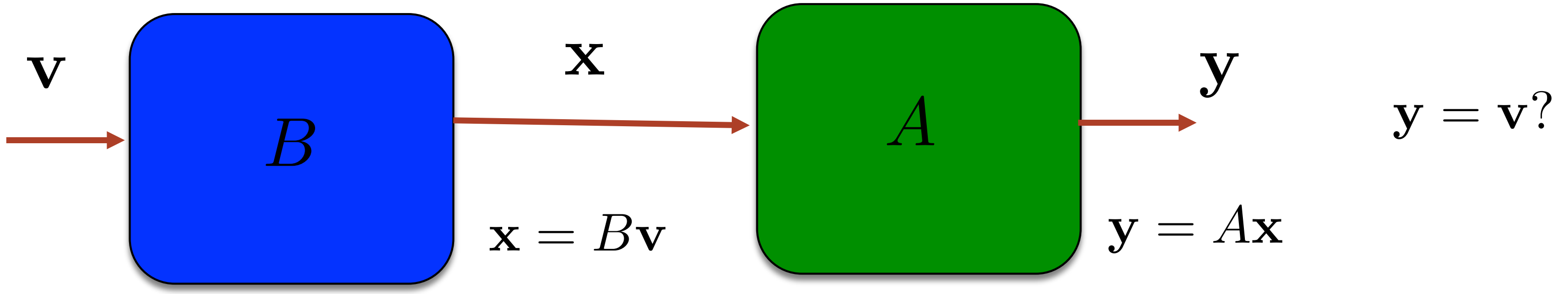
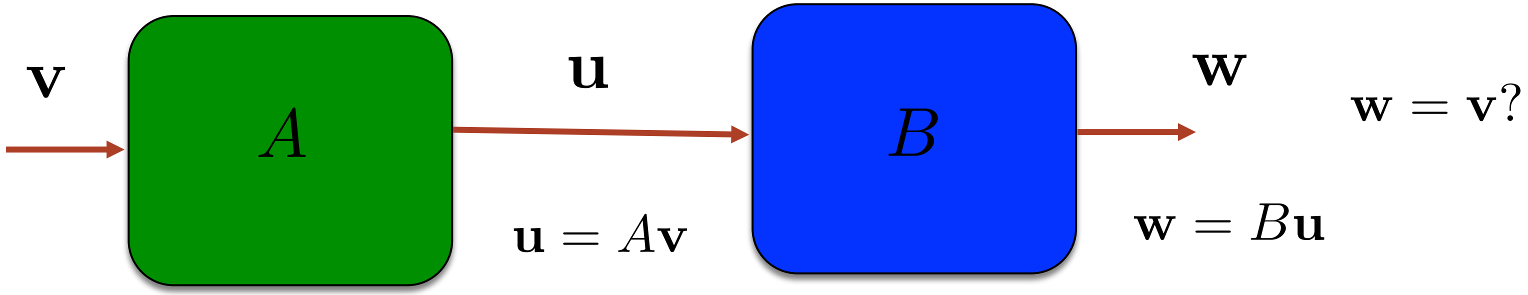
2.3 Invertibility and Elementary Matrices

From matrix **multiplication** to matrix **inverse**.

Let $\mathbf{v}, \mathbf{x}, \mathbf{y} \in \mathcal{R}^n$. Suppose A and B are $n \times n$ matrices.



Given an n x n matrix B, does there exist an n x n matrix A such that $\mathbf{y} = \mathbf{v}$ for any $\mathbf{v} \in \mathcal{R}^n$?



Furthermore, can this matrix also satisfy that $\mathbf{w} = \mathbf{v}$ for any $\mathbf{v} \in \mathcal{R}^n$?

Definitions

An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, B is called an **inverse** of A.

Properties:

1. For $n = 1$, the definition reduces to the multiplicative inverse ($ab = ba = 1$).
2. If B is an inverse of A, then A is an inverse of B, i.e., A and B are inverses to each other.

Example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \qquad B = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$$

$$\Rightarrow$$

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

$$BA = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Definitions

An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, B is called an **inverse** of A.

Question: Does each $n \times n$ matrix have an inverse?

Some example matrices that have no inverse:

1. $O \in \mathcal{M}_{n \times n}$.
2. $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$

since $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+2c & b+2d \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Definitions

An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, B is called an **inverse** of A.

Property:

If A is an invertible matrix, then A has exactly one inverse (denoted A^{-1}).

Proof

Definitions

An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, B is called an **inverse** of A.

Property:

If A is an invertible matrix, then A has exactly one inverse (denoted A^{-1}).

Theorem 2.2

Let A and B be $n \times n$ matrices.

(a) If A is invertible, then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

(b) If A and B are invertible, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

(c) If A is invertible, then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.

Proof

Definitions

An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, B is called an **inverse** of A.

Property:

If A is an invertible matrix, then A has exactly one inverse (denoted A^{-1}).

Theorem 2.2

Let A and B be $n \times n$ matrices.

(a) If A is invertible, then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

(b) If A and B are invertible, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

(c) If A is invertible, then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.

Corollary

Let $A_1, A_2, \ldots, A_k$ be $n \times n$ invertible matrix. Then the product $A_1 A_2 \cdots A_k$ is also invertible, and

$$(A_1 A_2 \cdots A_k)^{-1} = (A_k)^{-1}(A_{k-1})^{-1} \cdots (A_1)^{-1}$$

Proof

Symbolically, the inverse may be used to solve matrix equations:

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ A^{-1}(A\mathbf{x}) &= A^{-1}\mathbf{b} \\ (A^{-1}A)\mathbf{x} &= A^{-1}\mathbf{b} \\ I_n\mathbf{x} &= A^{-1}\mathbf{b} \\ \mathbf{x} &= A^{-1}\mathbf{b} \end{aligned}$$

$$\underbrace{\begin{aligned} x_1 + 2x_2 &= 4 \\ 3x_1 + 5x_2 &= 7 \end{aligned}}_{A\mathbf{x}=\mathbf{b}}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} -6 \\ 5 \end{bmatrix}$$

However, this method is computationally inefficient.

Elementary Matrices

- Every elementary row operation can be performed by matrix multiplication.
- For examples: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$
- 1. Multiplying row 2 of A by the scalar k:

$$\begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ kc & kd \end{bmatrix}$$

- 2. Interchanging rows 1 and 2 of A:

$$\begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

- 3. Adding k times row 1 to row 2:

$$\begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ ka + c & kb + d \end{bmatrix}$$

Definition

An $n \times n$ matrix E is called an **elementary matrix** if E can be obtained from I_n by a **single elementary row operation**.

$$\text{Example: } E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}.$$

Proposition

Let A be an $m \times n$ **matrix**, and let E be an $m \times m$ **elementary matrix** obtained by performing an **elementary row operation** on I_m. Then the product EA can be obtained from A by performing the identical elementary row operation on A.

$$\text{Example:} \quad A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \Rightarrow E_3 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 7 & 10 \end{bmatrix}.$$

Question: Is an elementary matrix always invertible?

Every elementary matrix is **invertible**. Furthermore, the inverse of an elementary matrix is also an elementary matrix.

Example: $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$.

Then,

$$E_1^{-1} = \begin{bmatrix} \quad & \quad & \quad \\ & & \\ & & \end{bmatrix},$$

$$E_2^{-1} = \begin{bmatrix} \quad & \quad & \quad \\ & & \\ & & \end{bmatrix},$$

$$E_3^{-1} = \begin{bmatrix} \quad & \quad & \quad \\ & & \\ & & \end{bmatrix}.$$

In general, if $E \in \mathcal{R}^{n \times n}$ is an elementary matrix that corresponds to some elementary row operation, then E^{-1} is the elementary matrix that corresponds to the reverse elementary row operation.

Think: How do we transform a matrix A to its reduced row echelon form by multiplying it with elementary matrices?

Theorem 2.3

Let A be an $m \times n$ matrix with reduced row echelon form R. Then there exists an invertible $m \times m$ matrix P such that $PA = R$.

Proof

Theorem 2.3

Let A be an $m \times n$ matrix with reduced row echelon form R. Then there exists an invertible $m \times m$ matrix P such that $PA = R$.

Corollary

The matrix equation $A\mathbf{x} = \mathbf{b}$ has the same solution set as $R\mathbf{x} = \mathbf{c}$, where $\begin{bmatrix} R & \mathbf{c} \end{bmatrix}$ is the reduced echelon form of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$.

Proof

Theorem E.1 (Column Correspondence Theorem)

Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots \mathbf{a}_n \end{bmatrix}$ be an $m \times n$ matrix whose reduced row echelon form is $R = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \cdots \mathbf{r}_n \end{bmatrix}$. Then

(a) If $\mathbf{a}_j$ is a linear combination of other columns of A, then $\mathbf{r}_j$ is a linear combination of the corresponding columns of R with the same coefficients.

(b) If $\mathbf{r}_j$ is a linear combination of other columns of R, then $\mathbf{a}_j$ is a linear combination of the corresponding columns of A with the same coefficients.

Example:

$$A = \begin{bmatrix} 1 & 2 & -1 & 2 & 1 & 2 \\ -1 & -2 & 1 & 2 & 3 & 6 \\ 2 & 4 & -3 & 2 & 0 & 3 \\ -3 & -6 & 2 & 0 & 3 & 9 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{matrix} \mathbf{a}_2 = 2\mathbf{a}_1 \\ \mathbf{a}_5 = -\mathbf{a}_1 + \mathbf{a}_4 \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} \mathbf{r}_2 = 2\mathbf{r}_1 \\ \mathbf{r}_5 = -\mathbf{r}_1 + \mathbf{r}_4 \end{matrix}$$

Theorem E.1

Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots \mathbf{a}_n \end{bmatrix}$ be an $m \times n$ matrix whose reduced row echelon form is $R = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \cdots \mathbf{r}_n \end{bmatrix}$. Then

(a) If $\mathbf{a}_j$ is a linear combination of other columns of A, then $\mathbf{r}_j$ is a linear combination of the corresponding columns of R with the same coefficients.

(b) If $\mathbf{r}_j$ is a linear combination of other columns of R, then $\mathbf{a}_j$ is a linear combination of the corresponding columns of A with the same coefficients.

****Proof*** (a) Suppose $\mathbf{a}_j = \sum_{k=1, k \neq j}^{n} c_k \mathbf{a}_k.$

You show part (b).

Properties of a Matrix in Reduced Row Echelon Form (in Appendix E)

Let R be an $m \times n$ matrix in reduced row echelon form. Then the following statements are true.

(a) A column of R is a pivot column if and only if it is nonzero and not a linear combination of the preceding columns of R.

(b) The jth pivot column of R is $\mathbf{e}_j$, the jth standard vector of $\mathcal{R}^m$, and hence the pivot columns of R are linearly independent.

(c) Every column of R is a linear combination of the pivot columns of R.

$$\begin{bmatrix} 1 & 0 & * & * & 0 & 0 & * \\ 0 & 1 & * & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

pivot columns

Example:

$$\begin{bmatrix} 1 & 0 & * & * & 0 & 0 & * \\ 0 & 1 & * & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

pivot columns

More precisely for Property (c):
Suppose that $\mathbf{r}_j$ is not a pivot column of R, and there are k pivot columns of R preceding it. Then $\mathbf{r}_j$ is a linear combination of the k preceding pivot columns, and the coefficients of the linear combination are the first k entries of $\mathbf{r}_j$. Furthermore, the other entries of $\mathbf{r}_j$ are zeros.

Example:

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ -1 & 1 & -8 & 1 \\ 2 & -1 & 13 & -2 \\ 1 & -1 & 8 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 5 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Properties of R.R.E.F Matrices
+ Theorem E.1 (Correspondence Theorem)
➔(1) Pivot columns of A are linearly independent.
(2) Each non-pivot column is a linear comb. of pivot columns.

pivot columns

Example:

$$A = \begin{bmatrix} 1 & 2 & -1 & 2 & 1 & 2 \\ -1 & -2 & 1 & 2 & 3 & 6 \\ 2 & 4 & -3 & 2 & 0 & 3 \\ -3 & -6 & 2 & 0 & 3 & 9 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

non-pivot columns

$$\mathbf{a}_2 = 2\mathbf{a}_1$$
$$\mathbf{a}_5 = -\mathbf{a}_1 + \mathbf{a}_4$$
$$\mathbf{a}_6 = -5\mathbf{a}_1 - 3\mathbf{a}_3 + 2\mathbf{a}_4$$

Theorem 2.4

The following statments are true for any matrix A.
(a) The pivot columns of A are linearly independent.
(b) Each nonpivot column of A is a linear combination of the previous pivot columns of A, where the coefficients of the linear combination are the entries of the corresponding column of the reduced row echelon form of A.

pivot columns

Example:

$$A = \begin{bmatrix} 1 & 2 & -1 & 2 & 1 & 2 \\ -1 & -2 & 1 & 2 & 3 & 6 \\ 2 & 4 & -3 & 2 & 0 & 3 \\ -3 & -6 & 2 & 0 & 3 & 9 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

non-pivot columns

$$\mathbf{a}_2 = 2\mathbf{a}_1$$
$$\mathbf{a}_5 = -\mathbf{a}_1 + \mathbf{a}_4$$
$$\mathbf{a}_6 = -5\mathbf{a}_1 - 3\mathbf{a}_3 + 2\mathbf{a}_4$$

Theorem 2.4

The following statments are true for any matrix A.
(a) The pivot columns of A are linearly independent.
(b) Each nonpivot column of A is a linear combination of the previous pivot columns of A, where the coefficients of the linear combination are the entries of the corresponding column of the reduced row echelon form of A.

Proof

Suppose R is the r.r.e.f. of A.

Then there exists invertible mxm matrix P such that

$R = PA$.

Let $\mathbf{r}_{p1} = \mathbf{e}_1$, $\mathbf{r}_{p2} = \mathbf{e}_2$, …, be pivot columns of R.

Then, $\mathbf{a}_{p1} = P\mathbf{e}_1$, $\mathbf{a}_{p2} = P\mathbf{e}_2$, …, are also l.i.

(b) Suppose $\mathbf{r}_j$ is a nonpivot column of R, then by property (c),

$\mathbf{r}_j = c_1\mathbf{r}_{p1} + c_2\mathbf{r}_{p2} + \ldots$

Then $\mathbf{a}_j = P^{-1}\,\mathbf{r}_j = P^{-1}\,(c_1\mathbf{r}_{p1} + \ldots) = c_1\,\mathbf{a}_{p1} + c_2\,\mathbf{a}_{p2} + \ldots$

*Uniqueness of the reduced row echelon form of a matrix:

Theorem 1.4 **(Introduced in Section 1.3 and partly proved in Section 1.4)**

Every matrix can be transformed into **one** and **only one** in **reduced echelon form** by means of a sequence of **elementary row operations**.

Proof Let $A \in \mathcal{R}^{m\times n}$ and R be a reduced row echelon form of A.

Claim I: Contents and positions of the pivot columns of R are uniquely determined by A.

Claim II: Contents and positions of the nonpivot columns of R are uniquely determined by A.

*Uniqueness of the reduced row echelon form of a matrix:

Theorem 1.4 **(Introduced in Section 1.3 and partly proved in Section 1.4)**

Every matrix can be transformed into **one** and **only one** in **reduced echelon form** by means of a sequence of **elementary row operations**.

Proof Let $A \in \mathcal{R}^{m\times n}$ and R be a reduced row echelon form of A.

Claim I: Contents and positions of the pivot columns of R are uniquely determined by A.

Proof

By Property (a) of R and the column correspondence property, we know that

$\mathbf{r}_i$ is a pivot column of R

$\Leftrightarrow$ $\mathbf{r}_i \neq \mathbf{0}$ and $\mathbf{r}_i$ is not a l.c. of the preceding columns of R (Property a)

$\Leftrightarrow$ $\mathbf{a}_i \neq \mathbf{0}$ and $\mathbf{a}_i$ is not a linear combination of the preceding columns of A.

Thus **positions** of the pivot columns of R are uniquely determined by positions of A's nonzero columns that are not linear combinations of the preceding columns.

By the Property (b) of R, the contents of the pivot columns of R are fixed ($\mathbf{e}_1$, $\mathbf{e}_2$, … in $\mathcal{R}^m$).

*Uniqueness of the reduced row echelon form of a matrix:

Theorem 1.4 **(Introduced in Section 1.3 and partly proved in Section 1.4)**

Every matrix can be transformed into **one** and **only one** in **reduced echelon form** by means of a sequence of **elementary row operations**.

Claim II: Contents and positions of the nonpivot columns of R are uniquely determined by A.

Proof

Suppose $\mathbf{r}_{n_1} = \mathbf{e}_1, \mathbf{r}_{n_2} = \mathbf{e}_2, \ldots, \mathbf{r}_{n_k} = \mathbf{e}_k$ are the pivot columns of R, and $\mathbf{r}_j$ is a non-pivot column of R.

$\Rightarrow$ $\mathbf{r}_{n_1}, \mathbf{r}_{n_2}, \ldots, \mathbf{r}_{n_k}$ are L.I. (property b) and $\mathbf{r}_j = c_1\mathbf{r}_{n_1} + c_2\mathbf{r}_{n_2} + \cdots + c_k\mathbf{r}_{n_k}$ for some $c_1, c_2, \ldots, c_k$. (property c)

$\Rightarrow$ By the column correspondence property, (actually, by Thm 2.4b)

$\mathbf{a}_{n_1}, \mathbf{a}_{n_2}, \ldots, \mathbf{a}_{n_k}$ are L.I. and

$$\mathbf{a}_j = c_1\mathbf{a}_{n_1} + c_2\mathbf{a}_{n_2} + \cdots + c_k\mathbf{a}_{n_k}.$$

Thus $c_1, c_2, \ldots, c_k$ are (and hence $\mathbf{r}_j$ is) uniquely determined by A.

Homework Set for 2.3

Section 2.3: Problems 1, 3, 5, 9, 11, 15, 17, 19, 21, 23, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 49, 51.