Equivalent Systems, Rank, and Row Space

Equivalent Systems and Row Equivalence of Matrices

Definition Two systems of m linear equations in n variables are **equivalent** if and only if they have exactly the same solution set.

Definition An (augmented) matrix $\mathbf{D}$ is **row equivalent** to a matrix $\mathbf{C}$ if and only if $\mathbf{D}$ is obtained from $\mathbf{C}$ by a finite number of row operations of types (I), (II), and (III).

Table 2.1 Row operations and their inverses

Type of Operation	Operation	Reverse Operation
(I)	$\langle i\rangle \leftarrow c\,\langle i\rangle$	$\langle i\rangle \leftarrow \frac{1}{c}\,\langle i\rangle$
(II)	$\langle j\rangle \leftarrow c\,\langle i\rangle + \langle j\rangle$	$\langle j\rangle \leftarrow -c\,\langle i\rangle + \langle j\rangle$
(III)	$\langle i\rangle \leftrightarrow \langle j\rangle$	$\langle i\rangle \leftrightarrow \langle j\rangle$

Theorem 2.2 If a matrix $\mathbf{D}$ is row equivalent to a matrix $\mathbf{C}$, then $\mathbf{C}$ is row equivalent to $\mathbf{D}$.

Theorem 2.3 Let $\mathbf{AX} = \mathbf{B}$ be a system of linear equations. If $[\mathbf{C}|\mathbf{D}]$ is row equivalent to $[\mathbf{A}|\mathbf{B}]$, then the system $\mathbf{CX} = \mathbf{D}$ is equivalent to $\mathbf{AX} = \mathbf{B}$.

Rank of a Matrix

Theorem 2.4 Every matrix is row equivalent to a unique matrix in reduced row echelon form.

Definition Let $\mathbf{A}$ be a matrix. Then the **rank** of $\mathbf{A}$ is the number of nonzero rows (that is, rows with nonzero pivot entries) in the unique reduced row echelon form matrix that is row equivalent to $\mathbf{A}$.

EXAMPLE: Since

$$A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 0 & 1 & -2 \\ 4 & 4 & -1 & -4 \\ 7 & 6 & 2 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 4 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we have

$$\text{rank } A = 3$$

Homogeneous Systems and Rank

DEFINITION: A system of linear equations is said to be **homogeneous** if it can be written in the form $A\mathbf{x} = \mathbf{0}$. Otherwise, it is **nonhomogeneous**.

Theorem 2.5 Let $\mathbf{AX} = \mathbf{O}$ be a homogeneous system in n variables.

(1) If rank$(\mathbf{A}) < n$, then the system has a nontrivial solution.

(2) If rank$(\mathbf{A}) = n$, then the system has only the trivial solution.

Proof. After the Gauss-Jordan method is applied to the augmented matrix $[\mathbf{A}|\mathbf{O}]$, the number of nonzero pivots equals rank$(\mathbf{A})$. Suppose rank$(\mathbf{A}) < n$. Then at least one of the columns is a nonpivot column, and so at least one of the n variables on the left side of $[\mathbf{A}|\mathbf{O}]$ is independent. Now, because this system is homogeneous, it is consistent. Therefore, the solution set is infinite, with particular solutions found by choosing arbitrary values for all independent variables and then solving for the dependent variables. Choosing a nonzero value for at least one independent variable yields a nontrivial solution.

On the other hand, suppose rank$(\mathbf{A}) = n$. Then, because $\mathbf{A}$ has n columns, every column on the left side of $[\mathbf{A}|\mathbf{O}]$ is a pivot column, and each variable must equal zero. Hence, in this case, $\mathbf{AX} = \mathbf{O}$ has only the trivial solution. □

Corollary 2.6 Let $\mathbf{AX} = \mathbf{O}$ be a homogeneous system of m linear equations in n variables. If $m < n$, then the system has a nontrivial solution.

Linear Combinations of Vectors

DEFINITION: The vector $\mathbf{y}$ defined by

$$\mathbf{y} = c_1\mathbf{v_1} + \ldots + c_p\mathbf{v_p}$$

where $\mathbf{v_1}, \ldots, \mathbf{v_p}$ are vectors and $c_1, \ldots, c_p$ are scalars, is called a **linear combination** of vectors $\mathbf{v_1}, \ldots, \mathbf{v_p}$.

EXAMPLE: Let $\mathbf{a_1} = [3, -3, 6]$, $\mathbf{a_2} = [5, -2, 1]$, and $\mathbf{a_3} = [-4, 4, -8]$ in $\mathbb{R}^3$. Consider the vector $\mathbf{v} = [7, -1, 4]$. Is $\mathbf{v}$ a linear combination of $\mathbf{a_1}, \mathbf{a_2}$, and $\mathbf{a_3}$?

Solution: We have

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 7 \\ 0 & 3 & 0 & 6 \\ 0 & -9 & 0 & -18 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 7 \\ 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{5}{3} & -\frac{4}{3} & \frac{7}{3} \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Since there is no row of the form $[0\ 0\ \ldots\ 0\ c]$, where c is a nonzero number, the corresponding system is consistent. Therefore $\mathbf{v}$ is a linear combination of $\mathbf{a_1}, \mathbf{a_2}$, and $\mathbf{a_3}$. Moreover, since the number of pivots (two) is lesser than the number of vectors (three), it follows that $\mathbf{v}$ can be written as a linear combination of $\mathbf{a_1}, \mathbf{a_2}$, and $\mathbf{a_3}$ in infinitely many ways.

The Row Space of a Matrix

Definition Let $\mathbf{A}$ be an $m \times n$ matrix. The subset of $\mathbb{R}^n$ consisting of all vectors that are linear combinations of the rows of $\mathbf{A}$ is called the **row space** of $\mathbf{A}$.

EXAMPLE: Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

The row space of A is the set of all linear combinations of the vectors

$$\mathbf{v}_1 = (1, 2, 3, 4)$$
$$\mathbf{v}_2 = (5, 6, 7, 8)$$
$$\mathbf{v}_3 = (0, 0, 1, 2)$$

Example

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & -2 \\ 4 & 0 & 1 \\ -2 & 4 & -3 \end{bmatrix}.$$

We want to determine whether $[5, 17, -20]$ is in the row space of $\mathbf{A}$. If so, $[5, 17, -20]$ can be expressed as a linear combination of the rows of $\mathbf{A}$, as follows:

$$[5, 17, -20] = c_1[3, 1, -2] + c_2[4, 0, 1] + c_3[-2, 4, -3].$$

Equating the coordinates on each side leads to the following system:

$$\begin{cases} 3c_1 + 4c_2 - 2c_3 = 5 \\ c_1 \quad\quad + 4c_3 = 17 \\ -2c_1 + c_2 - 3c_3 = -20 \end{cases}, \quad \text{whose matrix row reduces to} \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Hence, $c_1 = 5$, $c_2 = -1$, and $c_3 = 3$, and

$$[5, 17, -20] = 5[3, 1, -2] - 1[4, 0, 1] + 3[-2, 4, -3].$$

Therefore, $[5, 17, -20]$ is in the row space of $\mathbf{A}$. ■

Example

The vector $\mathbf{X} = [3, 5]$ is not in the row space of $\mathbf{B} = \begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix}$ because there is no way to express $[3, 5]$ as a linear combination of the rows $[2, -4]$ and $[-1, 2]$ of $\mathbf{B}$. That is, row reducing

$$\left[\mathbf{B}^T \middle| \mathbf{X}\right] = \left[\begin{array}{cc|c} 2 & -1 & 3 \\ -4 & 2 & 5 \end{array}\right] \quad \text{yields} \quad \left[\begin{array}{cc|c} 1 & -\frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 11 \end{array}\right],$$

thus showing that the corresponding linear system is inconsistent. ■

Row Equivalence Determines the Row Space

Lemma 2.7 Suppose that $\mathbf{x}$ is a linear combination of $\mathbf{q}_1, \ldots, \mathbf{q}_k$, and suppose also that each of $\mathbf{q}_1, \ldots, \mathbf{q}_k$ is itself a linear combination of $\mathbf{r}_1, \ldots, \mathbf{r}_l$. Then $\mathbf{x}$ is a linear combination of $\mathbf{r}_1, \ldots, \mathbf{r}_l$.

Proof. Because $\mathbf{x}$ is a linear combination of $\mathbf{q}_1, \ldots, \mathbf{q}_k$, we can write $\mathbf{x} = c_1\mathbf{q}_1 + c_2\mathbf{q}_2 + \cdots + c_k\mathbf{q}_k$ for some scalars $c_1, c_2, \ldots, c_k$. But, since each of $\mathbf{q}_1, \ldots, \mathbf{q}_k$ can be expressed as a linear combination of $\mathbf{r}_1, \ldots, \mathbf{r}_l$, there are scalars $d_{11}, \ldots, d_{kl}$ such that

$$\begin{cases} \mathbf{q}_1 = d_{11}\mathbf{r}_1 + d_{12}\mathbf{r}_2 + \cdots + d_{1l}\mathbf{r}_l \\ \mathbf{q}_2 = d_{21}\mathbf{r}_1 + d_{22}\mathbf{r}_2 + \cdots + d_{2l}\mathbf{r}_l \\ \quad\vdots \qquad \vdots \qquad\quad \vdots \qquad \ddots \qquad \vdots \\ \mathbf{q}_k = d_{k1}\mathbf{r}_1 + d_{k2}\mathbf{r}_2 + \cdots + d_{kl}\mathbf{r}_l \end{cases}.$$

Substituting these equations into the equation for $\mathbf{x}$, we obtain

$$\begin{aligned} \mathbf{x} = &\quad c_1(d_{11}\mathbf{r}_1 + d_{12}\mathbf{r}_2 + \cdots + d_{1l}\mathbf{r}_l) \\ &+ c_2(d_{21}\mathbf{r}_1 + d_{22}\mathbf{r}_2 + \cdots + d_{2l}\mathbf{r}_l) \\ &\qquad \vdots \qquad\qquad \vdots \qquad\quad \ddots \qquad \vdots \\ &+ c_k(d_{k1}\mathbf{r}_1 + d_{k2}\mathbf{r}_2 + \cdots + d_{kl}\mathbf{r}_l) \end{aligned}.$$

Collecting all $\mathbf{r}_1$ terms, all $\mathbf{r}_2$ terms, and so on, we get

$$\begin{aligned} \mathbf{x} = &\quad (c_1d_{11} + c_2d_{21} + \cdots + c_kd_{k1})\mathbf{r}_1 \\ &+ (c_1d_{12} + c_2d_{22} + \cdots + c_kd_{k2})\mathbf{r}_2 \\ &\qquad \vdots \qquad\qquad \vdots \qquad\quad \ddots \qquad \vdots \\ &+ (c_1d_{1l} + c_2d_{2l} + \cdots + c_kd_{kl})\mathbf{r}_l \end{aligned}.$$

Thus, $\mathbf{x}$ can be expressed as a linear combination of $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_l$. □

Theorem 2.8 Suppose that $\mathbf{A}$ and $\mathbf{B}$ are row equivalent matrices. Then the row space of $\mathbf{A}$ equals the row space of $\mathbf{B}$.

Proof. (Abridged) Let $\mathbf{A}$ and $\mathbf{B}$ be row equivalent $m \times n$ matrices. We will show that if $\mathbf{x}$ is a vector in the row space of $\mathbf{B}$, then $\mathbf{x}$ is in the row space of $\mathbf{A}$. (A similar argument can then be used to show that if $\mathbf{x}$ is in the row space of $\mathbf{A}$, then $\mathbf{x}$ is in the row space of $\mathbf{B}$.)

First consider the case in which $\mathbf{B}$ is obtained from $\mathbf{A}$ by performing a single row operation. In this case, the definition for each type of row operation implies that each row of $\mathbf{B}$ is a linear combination of the rows of $\mathbf{A}$ (see Exercise 19(a)). Now, suppose $\mathbf{x}$ is in the row space of $\mathbf{B}$. Then $\mathbf{x}$ is a linear combination of the rows of $\mathbf{B}$. But since each of the rows of $\mathbf{B}$ is a linear combination of the rows of $\mathbf{A}$, Lemma 2.7 indicates that $\mathbf{x}$ is in the row space of $\mathbf{A}$. By induction, this argument can be extended to the case where $\mathbf{B}$ is obtained from $\mathbf{A}$ by any (finite) sequence of row operations (see Exercise 20). □