9. Rank and Linear Dependence

Geometric Aspects of Matrices

Imagine a 3-dimension space marked by the *x*-*y*-*z* axis. A point in this space is identified by its coordinates (x_1, y_1, z_1). For the moment, we do not think of this is a row or column vector, just an ordered list of three values. Think of a vector as an arrow pointing from the origin to this point. Every vector (except the zero vector) has a direction and a length.

Imagine now two vectors (x_1, y_1, z_1) and (x_2, y_2, z_2) such that

- The vectors point in the same direction (they may have different lengths), or
- the vectors point in exactly the opposite direction, or
- one of the vectors is the zero vector.

Suppose $(x_1, y_1, z_1) \neq (0,0,0)$. Now consider linear combinations of the two vectors

$$a_1(x_1, y_1, z_1) + a_2(x_2, y_2, z_2).$$

The result is another vector that has either the same direction as, or the exact opposite direction to (x_1, y_1, z_1), or is the zero vector. If we consider the set of all possible $a_1(x_1, y_1, z_1) + a_2(x_2, y_2, z_2)$ combinations, we get a line going through the origin. Although the two vectors live in three-dimensional space, we say they **span** a single straight line. It is as though the set of all linear combinations live in a one-dimensional **subspace** within the three-dimensional space. Throughout, we will reserve the terms 'space' and 'subspace' only if it includes the zero vector, i.e., passes through the origin.

The two vectors share a certain property. If the two vectors point in the same or in exact opposite directions, then $(x_2, y_2, z_2) = c(x_1, y_1, z_1)$ for some *c*. The vectors satisfy the property:

$$c_1(x_1, y_1, z_1) + c_2(x_2, y_2, z_2) = 0 \text{ for some } c_1 \text{ and } c_2 \text{, not both equal to zero.} \tag{9.1}$$

We say that the two vectors (x_1, y_1, z_1) and (x_2, y_2, z_2) are **linearly dependent**.

Now suppose the two vectors (x_1, y_1, z_1) and (x_2, y_2, z_2) are non-zero, and point in different directions (i.e., neither in the same nor opposite directions). Then any linear combination will result in a third vector which could potentially be in a different direction from that of (x_1, y_1, z_1) and (x_2, y_2, z_2). If we consider the set of all possible $a_1(x_1, y_1, z_1) + a_2(x_2, y_2, z_2)$ linear combinations, we get a plane (a two-dimensional space) that intersects at the origin. The two vectors span a two-dimensional subspace within the three-dimensional space. These two vectors do not satisfy property (9.1). We say the two vectors are **independent**.

Is it possible for two vectors to span the whole three-dimensional space? No, two vectors can span at most a (two-dimensional) plane passing through the origin. Suppose we have a set of two independent vectors (x_1, y_1, z_1) and (x_2, y_2, z_2), and we add to the set a third vector (x_3, y_3, z_3). Will we be able to span the entire three-dimensional space? It depends. If (x_3, y_3, z_3) lies on the plane spanned by (x_1, y_1, z_1) and (x_2, y_2, z_2), or if (x_3, y_3, z_3) is the zero-vector, then no, the three vectors together will only span a two-dimensional space. Here the three vectors are dependent, i.e., they have the property that

$$c_1(x_1, y_1, z_1) + c_2(x_2, y_2, z_2) + c_3(x_3, y_3, z_3) = 0 \text{ for some } c_1, c_2, \text{ and } c_3 \text{, not all equal to zero.} \tag{9.2}$$

However, if (x_3, y_3, z_3) is not zero and does not lie on the plane spanned by (x_1, y_1, z_1) and (x_2, y_2, z_2), then yes, the three vectors together will span the entire three-dimensional space. These three vectors are independent, i.e., do not satisfy (9.2).

If we add yet another vector into our set of three independent vectors, then obviously the fourth vector must also lie in the three-dimensional space and can be written as a linear combination of the first three vectors. A set of four (or more) three-dimensional vectors can span at most three-dimensions. A set of four (or more) three-dimensional vectors must be dependent.

Although difficult to picture geometrically in higher dimensions, similar statements can be made of sets of n-dimensional vectors. If there are m such vectors, then the space spanned by those vectors will be at most $\min(m,n)$, and could be less.

Rank of a matrix Consider the matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}.$$

We can view this as a collection of column vectors, or a stack of row vectors.

$$\mathbf{A} = \begin{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} & \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} & \begin{bmatrix} a_{13} \\ a_{23} \\ \vdots \\ a_{m3} \end{bmatrix} & \cdots & \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} [a_{11} & a_{12} & a_{13} & \cdots & a_{1n}] \\ [a_{21} & a_{22} & a_{23} & \cdots & a_{2n}] \\ & & \vdots & & \\ [a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}] \end{bmatrix}.$$

We write these as

$$\mathbf{A} = [\mathbf{A}_1 \quad \mathbf{A}_2 \quad \mathbf{A}_3 \quad \cdots \quad \mathbf{A}_n] = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \mathbf{a}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_m^{\mathrm{T}} \end{bmatrix}$$

Suppose $m \le n$. Then the space spanned by $\mathbf{A}_1, \ldots, \mathbf{A}_n$ is at most m-dimensional subspace since these are m-vectors. It is called the **column space** of $\mathbf{A}$. The dimensionality of the column space of A is called the **column rank** of **A**. The space spanned by $\mathbf{a}_1^{\mathrm{T}}, \ldots, \mathbf{a}_m^{\mathrm{T}}$ is the **row space** of **A** and it is at most m-dimensional despite being n-vectors, because there are only m of these vectors. The dimension of the row space of **A** is called the **row rank** of **A**. Obviously, the column rank of an $(m \times n)$ matrix cannot exceed the number of columns n, and the row rank cannot exceed the number of rows m.

We can speak unambiguously of the **rank** of a matrix because the <u>column and row rank of any matrix **A** are always the same</u>. The following argument explains why: suppose the column rank of **A** is $r \le m$ where $r \le \min(m,n)$. This means you can find r independent column vectors in **A**. Collect these columns into a $(m \times r)$ matrix **C**. Every column in **A** can be written as a linear combination of the vectors in **C**, therefore we can write $\mathbf{A} = \mathbf{CR}$ for some $(r \times n)$ matrix **R**. This also says that every row of **A** is a linear combination of the rows of **R**. Since there are only r rows in **R**, it must be that $\text{row rank}(\mathbf{A}) \le r$, i.e.,

$$\text{row rank}(\mathbf{A}) \le \text{col rank}(\mathbf{A}) \tag{9.3}$$

Applying the same argument to the transpose $\mathbf{A}^{\mathrm{T}}$ we get

$$\text{row rank}(\mathbf{A}^{\mathrm{T}}) \le \text{col rank}(\mathbf{A}^{\mathrm{T}}).$$

Since row rank of a matrix is the column rank of its transpose, we have

$$\text{col rank}(\mathbf{A}) \le \text{row rank}(\mathbf{A}). \tag{9.4}$$

The two inequalities (9.3) and (9.4) gives

$$\text{col rank}(\mathbf{A}) = \text{row rank}(\mathbf{A}). \tag{9.5}$$

An $(m \times n)$ matrix $\mathbf{A}$ is **full-rank** if $rank(\mathbf{A}) = \min(m,n)$. E.g., the matrix

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

is full rank since its rank is 2 which is equal to $\min(2,3)$. The identity matrix

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

is full rank since any n-dimensional vector can be written as a linear combination of its columns (or rows).

The rank of reduced row echelon matrices The rank of row echelon matrices and reduced row echelon matrices are easy to determine – the rank is simply the number of pivots. Consider, for instance, the REF matrix

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

The pivot columns (1), (3), (4) are independent: there is no way to get (4) by combining (1) and (2) since the third elements of both are zero, and the third element of (4) is not. There is no way to get (2) by combining (1) and (3) since (3) would have to be set to zero. Similarly, there is no way to get (1) by combining (3) and (4). On the other hand, any of the non-pivot columns can be obtained by combining (1), (3) and (4). Similarly remarks can be made for the rows.

Row Echelon Forms of matrices are obtained by apply elementary row operators. The following result shows that we elementary row operations do not change the rank of a matrix. Therefore we can find the rank of a matrix by reducing it to its REF, then count the number of pivots.

<u>Result 1</u> Elementary row operations do not change the rank of a matrix.

Proof Suppose $\mathbf{B}$ is derived from $\mathbf{A}$ by applying a single elementary row operator. Denote the ith row of $\mathbf{B}$ by $\mathbf{b}_i$ and the jth row of $\mathbf{A}$ be $\mathbf{a}_j$, etc. Then either

- $\mathbf{b}_i = \mathbf{a}_j$ for some $i \neq j$ (result of a row swap) or
- $\mathbf{b}_i = c\mathbf{a}_i$ for some non-zero constant c, or
- $\mathbf{b}_i = \mathbf{a}_i + c\mathbf{a}_j$ for some $i \neq j$.

These are all linear combinations, so every row in $\mathbf{B}$ must reside in the row space of $\mathbf{A}$, therefore $rank(\mathbf{B}) \le rank(\mathbf{A})$. However, every row of $\mathbf{A}$ can be obtained from $\mathbf{B}$ by applying the opposite elementary row operation, so we also have $rank(\mathbf{A}) \le rank(\mathbf{B})$. In other words, $rank(\mathbf{A}) = rank(\mathbf{B})$

Geometric Aspects of Systems of Linear Equations

Suppose

$$\underbrace{\mathbf{A}}_{m\times n}\underbrace{\mathbf{x}}_{n\times 1} = \underbrace{\mathbf{b}}_{m\times 1} \tag{9.6}$$

is a system of m equations in n unknowns. Suppose $\mathbf{b} \neq \mathbf{0}$. The left hand side is a linear combination of the columns of **A**:

$$\mathbf{Ax} = \begin{bmatrix} \underbrace{\mathbf{A}_1}_{m\times 1} & \underbrace{\mathbf{A}_2}_{m\times 1} & \underbrace{\mathbf{A}_3}_{m\times 1} & \cdots & \underbrace{\mathbf{A}_n}_{m\times 1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + x_3\mathbf{A}_3 + \cdots + x_n\mathbf{A}_n .$$

The problem of solving (9.6) can be stated as: find all linear combinations of the columns of **A** that give **b**. This is only possible if **b** lies in the column space of **A**. If **b** does not lie in the column space of **A**, there is no solution. Otherwise, there could be one solution, or many solutions. If there are many solutions, the solution set could make up a line, or a plane, or a 3-dimensional hyperplane, etc. The solutions will be $(n\times 1)$ vectors.

In the case of a homogenous system of equations

$$\underbrace{\mathbf{A}}_{m\times n}\underbrace{\mathbf{x}}_{n\times 1} = \underbrace{\mathbf{0}}_{m\times 1} \tag{9.7}$$

The solutions will always exist. Since the solutions will always include $\mathbf{x} = \mathbf{0}$, the solutions will be sets of n-dimensional vectors that could be the zero matrix alone, or a one-dimensional subspace (a line passing through the origin), a two-dimensional subspace (a plane passing through the origin), etc. of $\mathbb{R}^n$. The solution subspace is called the null space of $\mathbf{A}$, $null(\mathbf{A})$. Recall that the solutions to systems of equations has dimension equal to the number of free parameters, which is n less the number of pivots. The rank of **A** is equal to the number of pivots. Therefore, the dimension of $null(\mathbf{A})$ is $n-r$ where r is the rank of $\mathbf{A}$.

Like the solutions to (9.7), the row space of **A** also resides in $\mathbb{R}^n$ (each of the m rows of **A** is an n-vector). The row space spans a space of dimension equal to its rank r. The dimension of the row space of **A** and the dimension of its null space adds up to n. The fact that every row in **A** and every solution **x** satisfies (9.7) means that every vector in the row space of **A** is orthogonal to every vector in the null space of **A**. If n is 2- or 3-, then every vector in the row space of **A** is perpendicular to every vector in the null space of **A**. We say that the row space of **A** and the null space of **A** are orthogonal complements.

We can apply the same discussion to the system

$$\underbrace{\mathbf{A}^{\mathrm{T}}}_{n\times m}\underbrace{\mathbf{x}}_{m\times 1} = \underbrace{\mathbf{0}}_{n\times 1} \tag{9.8}$$

where now the solution vectors make up a m dimensional subspace. The row space of $\mathbf{A}^{\mathrm{T}}$ is a space of dimension r. The solution space of (9.8), i.e., the null space of $\mathbf{A}^{\mathrm{T}}$, is a space of dimension $m-r$. The row space of $\mathbf{A}^{\mathrm{T}}$ is the column space of **A**. The vectors in the two spaces satisfy (9.8). Therefore the column space of **A** and the null space of $\mathbf{A}^{\mathrm{T}}$ are orthogonal complements. This result and the result in the previous paragraph make up the "Fundamental Theorem of Linear Algebra".

Further properties of the rank of a matrix

<u>Result 2</u> For any matrices **A** and **B** for which **AB** exists, we have

$$rank(\mathbf{AB}) \le \min(rank(\mathbf{A}), rank(\mathbf{B})) \,. \tag{9.9}$$

Proof Suppose $\mathbf{A}$ is $(m \times p)$ and $\mathbf{B}$ is $(p \times n)$ so that $\mathbf{AB}$ is $(m \times n)$. The columns of the matrix product $\mathbf{AB}$ can be viewed as a collection of *n* linear combinations of the columns of **A**, therefore $rank(\mathbf{AB}) \le rank(\mathbf{A})$. The rows of $\mathbf{AB}$ can be viewed as a collection of m linear combinations of the rows of **B**, therefore $rank(\mathbf{AB}) \le rank(\mathbf{B})$. This proves (9.9).

<u>Result 3</u> A square matrix **A** has an inverse if and only if it is full rank.

Proof Suppose we can reduce a square matrix **A** to the identity matrix using elementary row operations (i.e., the inverse exists). Since elementary row operations do not change the rank of a matrix, **A** must have the same rank as the identity matrix. Since the identity matrix has full rank, **A** has full rank.

If **A** has full rank, then its reduced echelon form must be full rank since elementary row operators do not change the rank of a matrix. The only full rank reduced echelon form is the identity matrix. Therefore $\mathbf{A}^{-1}$ exists.

<u>Result 4</u> If $\mathbf{A}$ is a full rank $(m \times m)$ matrix, and **B** is an $(m \times n)$ matrix of rank *r*, then $rank(\mathbf{AB}) = r$.

Proof Suppose $\mathbf{E}_n \mathbf{E}_{n-1} \cdots \mathbf{E}_1$ are the sequence of elementary row operations that reduce **A** to $\mathbf{A}^{-1}$. Applying these to the matrix **AB** reduces **AB** to **B** but does not change the rank, which proves the result.

Exercises

1. Show that $rank(\mathbf{A}^{\mathrm{T}}\mathbf{A}) = rank(\mathbf{A})$ where $\mathbf{A}$ is $n \times k$.

Ans: The row spaces of both $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ are subspaces of $\mathbb{R}^k$. Furthermore, both $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ have the same null space: if **x** satisfies $\mathbf{Ax} = \mathbf{0}$, then **x** satisfies $\mathbf{A}^{\mathrm{T}}\mathbf{Ax} = \mathbf{0}$; if $\mathbf{A}^{\mathrm{T}}\mathbf{Ax} = \mathbf{0}$, then $\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{Ax} = (\mathbf{Ax})^{\mathrm{T}}\mathbf{Ax} = \mathbf{0}$, which implies that $\mathbf{Ax} = \mathbf{0}$, since $(\mathbf{Ax})^{\mathrm{T}}\mathbf{Ax}$ is a sum of squares. Both $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ must therefore have the same rank since their ranks are equal to *k* minus the dimension of the respective null spaces.