1 Vector

Before starting discussions on the next part (Rank) of this topic, we would need to develop some fundamental tools. Vectors will be an indispensable tool.

Given a natural number d (that is a), we take $\mathbb{R}^d$ to denote a d-dimensional vector and its corresponding space in which each dimension has a domain of (Can you recall the meaning of domain and range?). Vector, in general, is defined as either a $d \times 1$ matrix (column vector) or $1 \times d$ (row vector). Therefore, a d-dimensional vector is of the form $[v_1, v_2, \cdots, v_d]$ in which $v_i, \forall i \in [1,2, \dots, d]$ is a real value. To represent a vector in written form, we take boldface to denote a vector e.g. $\boldsymbol{v} = [v_1, v_2, \cdots, v_d]$.

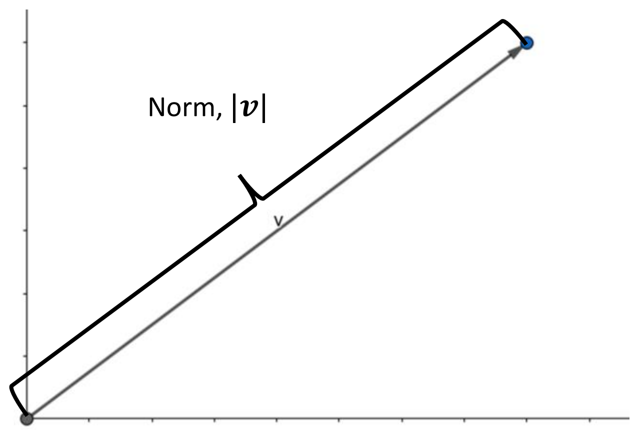

Figure 1 Vector and its Norm

In particular, we name a vector to be zero vector, $\mathbf{0}$, when all $v_i = 0, \forall i \in [1,2, \dots, d]$ or, for brevity, $\mathbf{0} = [0,0, \cdots ,0]$. When it comes to the length of a vector $\boldsymbol{v} = [v_1, v_2, \cdots, v_d]$, we make use of the term *norm* or symbolically $|\boldsymbol{v}|$ such that

$$|\boldsymbol{v}| = \sqrt{{v_1}^2 + {v_2}^2 + \cdots + {v_d}^2} = \sqrt{\sum_{\boldsymbol{i=1}}^{\boldsymbol{d}} {v_i}^2}$$

Obviously, for zero vector, the norm is zero.

$$|\boldsymbol{o}| = \sqrt{0^2 + 0^2 + \cdots + 0^2} = \sqrt{\sum_{\boldsymbol{i=1}}^{\boldsymbol{d}} 0^2} = 0$$

On the other hand, we refer a vector to be *unit vector* if $|\boldsymbol{v}| = 1$.

Example 1:

Consider a 3-dimensional vector $\boldsymbol{v} = [\frac{1}{3}, \frac{2}{3}, \frac{2}{3}]$, its norm is computed as follows

$$|\boldsymbol{v}| = \sqrt{(\frac{1}{3})^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2} = 1$$

And, we say that vector $\boldsymbol{v}$ is an unit vector.

2 Linear Independence

Suppose there exists a group of *m* vectors $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_m}$ having the . We define *Linear Combination* of those *m* vectors as

$$\sum_{i=1}^{d} c_i \, \boldsymbol{u_i}$$

given real value c_i $\forall i \in [1,2, \dots, m]$.

Example 2:

Consider three vectors, $\boldsymbol{u_1} = [6,7,8]$, $\boldsymbol{u_2} = [3,2,1]$ & $\boldsymbol{u_3} = [6,5,4]$. Clearly, these three vectors are in 3-dimensional. We can check that the vector $\boldsymbol{u}$ can be re-written as the following linear combination.

$$\boldsymbol{u_1} = [6,7,8] = 3 \cdot [6,5,4] + (-4) \cdot [3,2,1] = 3 \cdot \boldsymbol{u_2} + (-4) \cdot \boldsymbol{u_3}$$

Based on the result, we say that vector $\boldsymbol{u}$ is a linear combination of the vectors $\boldsymbol{u_2}$ and $\boldsymbol{u_3}$ in $\mathbb{R}^3$.

Check point: Could $\boldsymbol{u_2}$ be linear combination of the vector $\boldsymbol{u_1}$ and $\boldsymbol{u_3}$? How about the linear combination of $\boldsymbol{u_3}$?

Definition 1:
Let $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_m}$ be vectors of the same dimension and $c_1, c_2, \cdots, c_m$ be real numbers (or constants). We say that $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_m}$ are linearly dependent if the following conditions are satisfied together.

1. $$\sum_{i=1}^{d} c_i\, u_i = 0$$

2. $c_1, c_2, \cdots, c_m$ are not .

Otherwise, the m vectors $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_m}$ are linearly independent.

Example 3:

Consider 3 vectors $\boldsymbol{u_1} = [1,2], \boldsymbol{u_2} = [0,1]$ & $\boldsymbol{u_3} = [3,4]$. We would like to check with the dependence among these vectors. We can find that

$$3 \cdot \boldsymbol{u_1} + (-2) \cdot \boldsymbol{u_2} + (-1) \cdot \boldsymbol{u_3} = 3 \cdot [1,2] + (-2) \cdot [0,1] + (-1) \cdot [3,4] = [0,0] = \boldsymbol{0}$$

Apparently, these 3 vectors are linearly dependent by Definition 1. By the same token, we can check the dependence for $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$. Consider

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} = \boldsymbol{0}$$

And we seek to find out the possible values of c_1 and c_2. And, we find that the only possible values for the two constants are , then we conclude that $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$ are linearly .

Check point: How about the linear dependencies for (1) $\boldsymbol{u_1}$ and $\boldsymbol{u_3}$ and (2) $\boldsymbol{u_2}$ and $\boldsymbol{u_3}$?

3 Rank of a Matrix

Definition 2:

The rank of a matrix A, denoted as $r(A)$, is the maximum (largest) number of linearly independent row or column vector of A.

Rank, a real number, is a vital element associated with any matrix.

Example 4:

Consider a matrix $A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 5 \\ 2 & -1 & 4 \end{pmatrix}$. For the first two columns, they are proportional to each other (first column is multiple of second column by a factor of -2). Hence, there are at most 2 linearly independent column and therefore $r(A) \leq 2$.

Example 5:

Consider a matrix $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}$. From Example 3, as a matter of fact, $A = (u_1^T\ u_2^T\ u_3^T)$. Now, we have obviously seen that those 3 rows of A are linearly dependent. On the other hand, the example has shown that vectors $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$ [or u_1^T & u_2^T correspondingly] are linearly independent. Following the Definition 2, we conclude that the maximum number of column vector of A that are linearly independent is 2. Therefore, we have $r(A) = 2$.

Similarly, when you extend to considering 1) $\boldsymbol{u_1}$ and $\boldsymbol{u_3}$ and (2) $\boldsymbol{u_2}$ and $\boldsymbol{u_3}$, it will also lead to the same conclusion.

Recall that we learnt elementary row operations in previous topic. Suppose there are two matrices A and B, both of them are said to be *row-equivalent* if we can transform A to B, or vice versa by applying elementary row operations.

Note that applying elementary row operations on matrix do **NOT** change its rank!

Theorem 1:

If matrices A and B are row equivalent, they have the same rank.

Example 6:

Based on Example 5, consider the matrix $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}$ has $r(A) = 2$. We

find that the subsequent matrices are row-equivalent. They are

$\begin{pmatrix} 1 & 2 \\ 0 & 3 \\ 3 & 4 \end{pmatrix}$, $\begin{pmatrix} 1 & 5 \\ 0 & 3 \\ 3 & 4 \end{pmatrix}$ and $\begin{pmatrix} 1 & 5 \\ 0 & 3 \\ 3 & 1 \end{pmatrix}$ [Check how to arrive at this conclusion of row-equivalent!]

By Theorem 1, we can claim that the concerned matrices all have rank equal to 2.

Other than elementary row operations, we also discussed the importance of (reduced) row echelon form. We have the following theorem.

Theorem 2:

If matrix A is of row echelon form, rank of A, $r(A)$, equals the number of non-zero rows of A.

Example 7:
Consider the following matrices, the rank of each corresponding matrix:

$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$ Rank: $r(A) = 2$

$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ Rank: $r(A) = 3$

$\begin{pmatrix} 1 & 2 & 3 & 2 \\ 0 & -1 & -1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ Rank: $r(A) = 2$

Based on Theorems 1 and 2, they provide us with a handy way to compute the rank of a matrix A.

i. Convert given matrix A to row echelon form, say A' by elementary row operations.
ii. Count the total number of   of A'.

Example 8:

Revisit the Example 4 where matrix $A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 5 \\ 2 & -1 & 4 \end{pmatrix}$. We intentionally omit the steps on elementary row operations and just show you below the results after the operations. (We leave the intermediate steps to you as an exercise.)

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 5 \\ 2 & -1 & 4 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

We can see that there are 2 linearly independent row vectors and therefore $r(A) = 2$ which agrees with the conclusion of Example 4.

Example 9:

Consider the matrix $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}$. Similar to Example 8, we carry out elementary row operations. We can see that

$$B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Clearly, the rank of matrix B is 2, i.e, $r(B) = 2$.

In fact, matrix B is actually the transpose of the given matrix A of Example 5. In other words,

$$B = A^T \text{ or } A = B^T$$

The results of rank in these 2 examples are the same. It naturally lead us to questioning whether or not there is underlying relationship between rank and transpose of a matrix. We have the following theorem for this observation.

Theorem 3:

The rank of a matrix is equal to the rank of transpose of the matrix.

Theorem 4:

Suppose that there is a matrix A such that $\boldsymbol{r_i}$ denotes the i^{th} row vectors of A and $\boldsymbol{r_1}, \boldsymbol{r_2}, \dots$ & $\boldsymbol{r_k}$ are linearly independent but $\boldsymbol{r_1}, \boldsymbol{r_2}, \dots, \boldsymbol{r_k}$ & $\boldsymbol{r_{k+1}}$ are linearly dependent. Then $\boldsymbol{r_{k+1}}$ must be a linearly combination of $\boldsymbol{r_1}, \boldsymbol{r_2}, \dots$ & $\boldsymbol{r_k}$.

Proof:

Given that $\boldsymbol{r_1}, \boldsymbol{r_2}, \dots, \boldsymbol{r_k}$ & $\boldsymbol{r_{k+1}}$ are linearly dependent, $\exists$ $c_1, c_2, \dots, c_k$ & c_{k+1} are real constants s.t.

i. $c_1\boldsymbol{r_1} + c_2\boldsymbol{r_2} + \dots + c_k\boldsymbol{r_k} + c_{k+1}\boldsymbol{r_{k+1}} = 0$

ii. $c_1, c_2, \dots, c_k$ & c_{k+1} are not all zero.

It is known that c_{k+1} cannot be zero. If $c_{k+1} = 0$, by (i),

$$\begin{aligned} & c_1\boldsymbol{r_1} + c_2\boldsymbol{r_2} + \dots + c_k\boldsymbol{r_k} + c_{k+1}\boldsymbol{r_{k+1}} \\ &= c_1\boldsymbol{r_1} + c_2\boldsymbol{r_2} + \dots + c_k\boldsymbol{r_k} \\ &= 0 \end{aligned}$$

It implies that $c_1, c_2, \dots, c_k$ are not all zero (or at least one of them is non-zero) and $\boldsymbol{r_1}, \boldsymbol{r_2}, \dots$ & $\boldsymbol{r_k}$ becomes linearly dependent which leads to a contradiction to the fact that $\boldsymbol{r_1}, \boldsymbol{r_2}, \dots$ & $\boldsymbol{r_k}$ are linearly independent.

Now, $c_{k+1} \neq 0$ and, by (i), we have

$$-c_{k+1}\boldsymbol{r_{k+1}} = c_1\boldsymbol{r_1} + c_2\boldsymbol{r_2} + \cdots + c_k\boldsymbol{r_k}$$

$$\boldsymbol{r_{k+1}} = (-\frac{c_1}{c_{k+1}})\boldsymbol{r_1} + (-\frac{c_2}{c_{k+1}})\boldsymbol{r_2} + \cdots + (-\frac{c_k}{c_{k+1}})\boldsymbol{r_k}$$

Therefore, $\boldsymbol{r_{k+1}}$ is a linear combination of $\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots$ & $\boldsymbol{r_k}$.

Physical meaning of Theorem 4:

If a matrix has rank k, there exists exactly k "effective" rows. Except these k rows, each other rows can be expressed as a linear combination of those k rows. See the next example for more information.

Example 10:

Consider a matrix $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 3 & 2 & 1 & 0 \end{pmatrix}$. We can verify that $r(A) = 2$ by elementary row operations (the intermediate steps are left as an exercise.). We note that the third (last) row can be expressed as follows:

$$[\,3, 2, 1, 0] = [5, 6, 7, 8] + (-2) \cdot [1, 2, 3, 4]$$

which agrees with the Theorem 4.