Lecture Notes: Rank of a Matrix

Yufei Tao
Department of Computer Science and Engineering
Chinese University of Hong Kong
taoyf@cse.cuhk.edu.hk

1 Linear Independence

Definition 1. *Let $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_m}$ be vectors of the same dimensionality. Given real values $c_1, ..., c_n$,*

$$\sum_{i=1}^{n} c_i \, \boldsymbol{r_i}$$

is called a **linear combination** *of $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_m}$.*

Definition 2. *Let $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_m}$ be vectors of the same dimensionality. If we can find real values $c_1, ..., c_n$ such that*

- *$c_1, ..., c_n$ are not all zero;*
- $\sum_{i=1}^{n} c_i \, \boldsymbol{r_i} = \mathbf{0}$

then we say that $\boldsymbol{r_1}$, ..., $\boldsymbol{r_m}$ are **linearly dependent**. *Otherwise, $\boldsymbol{r_1}$, ..., $\boldsymbol{r_m}$ are* **linearly independent**.

Example 1. Consider vectors

$$\begin{aligned} \boldsymbol{r_1} &= [1, 2] \\ \boldsymbol{r_2} &= [0, 1] \\ \boldsymbol{r_3} &= [3, 4]. \end{aligned}$$

$\boldsymbol{r_1}$, $\boldsymbol{r_2}$, and $\boldsymbol{r_3}$ are linearly dependent because

$$3\boldsymbol{r_1} - 2\boldsymbol{r_2} - \boldsymbol{r_3} = \mathbf{0}.$$

On the other hand, $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are linearly independent because (as you can verify easily) the following equation has a unique solution $c_1 = c_2 = 0$:

$$c_1\boldsymbol{r_1} + c_2\boldsymbol{r_2} = \mathbf{0}. \tag{1}$$

□

Lemma 1. *Suppose that $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_m}$ are linearly independent, but $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_{m+1}}$ are linearly dependent. Then, $\boldsymbol{r_{m+1}}$ must be a linear combination of $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_m}$.*

Proof. Since $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_{m+1}}$ are linearly dependent, there exist $c_1, ..., c_{m+1}$ such that (i) they are not all zero, and (ii)

$$c_1\boldsymbol{r_1} + c_2\boldsymbol{r_2} + ... + c_m\boldsymbol{r_m} + c_{m+1}\boldsymbol{r_{m+1}} \quad = \quad 0.$$

Note that c_{m+1} cannot be 0. Otherwise, it will follow that $c_1\boldsymbol{r_1} + c_2\boldsymbol{r_2} + ... + c_m\boldsymbol{r_m} = 0$. Since $c_1, ..., c_m$ cannot be all zero, this means that $\boldsymbol{r_1}, ..., \boldsymbol{r_m}$ were linearly dependent, which is a contradiction.

Now that $c_{m+1} \neq 0$, we have:

$$\boldsymbol{r_{m+1}} \quad = \quad \frac{c_1}{c_{m+1}}\boldsymbol{r_1} + \frac{c_2}{c_{m+1}}\boldsymbol{r_2} + ... + \frac{c_m}{c_{m+1}}\boldsymbol{r_m}.$$

Therefore, $\boldsymbol{r_{m+1}}$ is a linear combination of $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_m}$. □

For instance, in Example 1, we know that $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, and $\boldsymbol{r_3}$ are linearly dependent, but $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are linearly independent. Hence, $\boldsymbol{r_3}$ must be a linear combination of $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$. Indeed, $\boldsymbol{r_3} = 3\boldsymbol{r_1} - 2\boldsymbol{r_2}$.

2 Rank of a Matrix

Definition 3. *The* **rank** *of a matrix* $\boldsymbol{A}$*—denoted as* $\text{rank}\,\boldsymbol{A}$*—is the maximum number of linearly independent row vectors of* $\boldsymbol{A}$*.*

Example 2. Consider matrix

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

Let $\boldsymbol{r_i}$ ($i \in [1, 3]$) be the i-th row of $\boldsymbol{A}$. The rank of $\boldsymbol{A}$ cannot be 3 because we have seen from Example 1 that $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, and $\boldsymbol{r_3}$ are linearly dependent. On the other hand, we already know that $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are linearly independent. It follows that 2 is the maximum number of row vectors that are linearly independent. Therefore, $\text{rank}\,\boldsymbol{A} = 2$. □

The above example shows a method for calculating the rank of a matrix. However, the method is not easy to apply when the matrix is large in dimensions. Next, we will give an alternative method for rank computation which is much easier to use.

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ both be $m \times n$ matrices. Recall that $\boldsymbol{A}$ and $\boldsymbol{B}$ are said to be *row-equivalent* if we can convert $\boldsymbol{A}$ to $\boldsymbol{B}$ by applying the following *elementary row operations*:

1. Switch two rows of $\boldsymbol{A}$.
2. Multiply all numbers of a row of $\boldsymbol{A}$ by the same non-zero value.
3. Let $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ be two row vectors of $\boldsymbol{A}$. Update row $\boldsymbol{r}_i$ to $\boldsymbol{r}_i + c \cdot \boldsymbol{r}_j$, where c can be any real value.

Next, we refer to the above as Operation 1, 2, and 3, respectively.

Lemma 2. *If* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are row-equivalent, then they have the same rank.*

Proof. See appendix. □

Example 3. We already know from Example 2 that the following matrix has rank 2.

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{bmatrix}.$$

By Lemma 2, we know that all the following matrices also have rank 2:

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 5 \\ 0 & 3 \\ 3 & 4 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 & 5 \\ 0 & 3 \\ 3 & 1 \end{bmatrix}$$

□

We say that a row of a matrix is a *non-zero row* if the row contains at least one non-zero value. Then, we have the following fact:

Lemma 3. *If matrix* $\boldsymbol{A}$ *is in row echelon form, then* $\operatorname{rank} \boldsymbol{A}$ *is the number non-zero rows of* $\boldsymbol{A}$.

We omit a proof for the lemma, which is a good exercise for you.

Example 4. The ranks of the following matrices are 2 and 3, respectively.

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

□

Lemmas 2 and 3 suggest the following approach to compute the rank of a matrix $\boldsymbol{A}$. First, convert $\boldsymbol{A}$ to a matrix $\boldsymbol{A}'$ of row echelon form, and then, count the number of non-zero rows of $\boldsymbol{A}'$.

Example 5. Next, we use the approach to calculate the rank of the matrix in Example 2 (in the derivation below, $\Rightarrow$ indicates applying row elementary operations):

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

□

Example 6. Compute the rank of the following matrix:

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

Solution.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 3 & 2 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -4 & -8 & -12 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence, the original matrix has rank 2. □

Lemma 4. *Suppose that $\boldsymbol{A}$ has rank k, and that $\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_k}$ are k linearly independent row vectors of $\boldsymbol{A}$. Let $\boldsymbol{r}$ be any row vector of $\boldsymbol{A}$. It must hold that $\boldsymbol{r}$ is a linear combination of $\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_k}$.*

Proof. First, suppose that $\boldsymbol{r}$ is not equal to any of $\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_k}$. By definition of rank, we know that $\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_k}$, and $\boldsymbol{r}$ must be linearly dependent. Then, the lemma follows from Lemma 1. Now consider that $\boldsymbol{r}$ is one of $\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_k}$. Without loss of generality, suppose that $\boldsymbol{r} = \boldsymbol{r_1}$. Then, we have $\boldsymbol{r_1} = 1\,\boldsymbol{r_1} + 0\,\boldsymbol{r_2} + ... + 0\,\boldsymbol{r_k}$. □

For instance, consider the matrix in Example 6. We know that the first two rows are linearly independent, but the three rows together are linearly dependent. Lemma 1 indicates that the 3rd row must be a linear combination of the first two rows. This is true because:

$$\begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix} = (-2) \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix}$$

3 An Important Property of Ranks

In this section, we will prove a non-trivial lemma about ranks.

Lemma 5. *The rank of a matrix $\boldsymbol{A}$ is the same as the rank of $\boldsymbol{A}^T$.*

Proof. Define the *column-rank* of $\boldsymbol{A}$ to be the maximum number of independent column vectors of $\boldsymbol{A}$. Note that the column-rank of $\boldsymbol{A}$ is exactly the same as the rank of $\boldsymbol{A}^T$. Hence, to prove the lemma, it suffices to show that the rank of $\boldsymbol{A}$ is the same as the column-rank of $\boldsymbol{A}$.

We first show:

- If $\boldsymbol{A}$ is in row echelon form, then the rank of $\boldsymbol{A}$ is at most its column-rank.
- Elementary row operations on $\boldsymbol{A}$ do not change its column rank.

The proofs of the above facts are left to you as an exercise. Since elementary row operations on $\boldsymbol{A}$ do not change its rank, combining both facts shows that the rank of $\boldsymbol{A}$ is at most its column-rank.

On the other hand, reversing the above argument shows that the column-rank of $\boldsymbol{A}$ is at most its rank. With this, we complete the lemma. □

Example 7. Let

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

Compute the rank of A^T.

Solution. From Example 6, we know that the rank of A is 2. Lemma 5 tells us that the rank of A^T must also be 2. □

Appendix

Proof of Lemma 2

Let us assume that $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ after applying *one* elementary row operation. To prove the lemma, it suffices to show that $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same rank. Let $\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_m}$ be the row vectors of $\boldsymbol{A}$.

- *Case 1: Operation 1 was applied.* In this case, $\boldsymbol{B}$ has exactly the same row vectors as $\boldsymbol{A}$. Hence, they have the same rank.

- *Case 2: Operation 2 was applied.* We will show that if a set of rows in $\boldsymbol{A}$ is linearly dependent (or independent), then the corresponding[1] rows in $\boldsymbol{B}$ are also linearly dependent (or independent, resp.). This implies that $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same rank.

  Without loss of generality, assume that the row operation multiplies the first row $\boldsymbol{r_1}$ of $\boldsymbol{A}$ with a value c. In other words, the row vectors of $\boldsymbol{B}$ are $c\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_m}$. Let R be a set of rows in $\boldsymbol{A}$. If $\boldsymbol{r_1} \notin R$, then all the rows of R remain the same in B. Hence, whether the rows of R are linearly independent remains the same in $\boldsymbol{B}$.

  It remains to consider that $\boldsymbol{r_1} \in R$. Define $k = |R|$. Without loss of generality, assume that $R = \{\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_k}\}$. The set of corresponding rows in B is $R' = \{c\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_k}\}$. If there exist $\alpha_1, ..., \alpha_k$ such that (i) they are not all zero, and (ii)

  $$\alpha_1 \boldsymbol{r_1} + \alpha_2 \boldsymbol{r_2} + ... + \alpha_k \boldsymbol{r_k} \quad = \quad \mathbf{0}$$

  it holds that

  $$\frac{\alpha_1}{c}(c\boldsymbol{r_1}) + \alpha_2 \boldsymbol{r_2} + ... + \alpha_k \boldsymbol{r_k} \quad = \quad \mathbf{0}$$

  In other words, we can find $\beta_1, ..., \beta_k$ such that (i) they are not all zero, and (ii)

  $$\beta_1 \boldsymbol{r_1} + \beta_2 \boldsymbol{r_2} + ... + \beta_k \boldsymbol{r_k} \quad = \quad \mathbf{0}.$$

  Hence, that R being linearly dependent implies R' being linearly dependent. The reverse of the above argument shows that R' being linearly dependent implies R being linearly dependent.

- *Case 3: Operation 3 was applied.* The proof of this case is similar to the proof of Case 2, and is left to you as an exercise.

[1]The correspondence here means: if we take the i-th row of $\boldsymbol{A}$, we also take the i-th row of $\boldsymbol{B}$.