

## 1.6 A Review of Matrix Theory

### Notes

In this section admit general matrices:  $A$  is  $m \times n$  of rank  $r$ .

Definition: The rank of  $A$  is the number,  $r$ , of linearly independent columns in  $A$ .

IN  $Ax$  is always a combination of the columns of  $A$ ; it is in the column space of  $A$ , denoted  $\mathcal{R}(A)$ . The system  $Ax=b$  has a solution exactly when  $b$  is also in the column space of  $A$ .

Definition: The nullspace of  $A$ , denoted  $\mathcal{N}(A)$ , is the set of vectors  $x$  satisfying  $Ax=0$ .

For  $A \in \mathbb{R}^{m \times n}$ :

The column space  $\mathcal{R}(A)$  is a subset of  $\mathbb{R}^m$  and has dimension  $r$ .  
The nullspace  $\mathcal{N}(A)$  is a subset of  $\mathbb{R}^n$  and has dimension  $n-r$ .

10 If the matrix  $A$  has linearly independent columns then:

- (1) The nullspace contains only the point  $x=0$ .
- (2) The solution to  $Ax=b$  is unique (if it exists).
- (3) The rank is  $r=n$ .

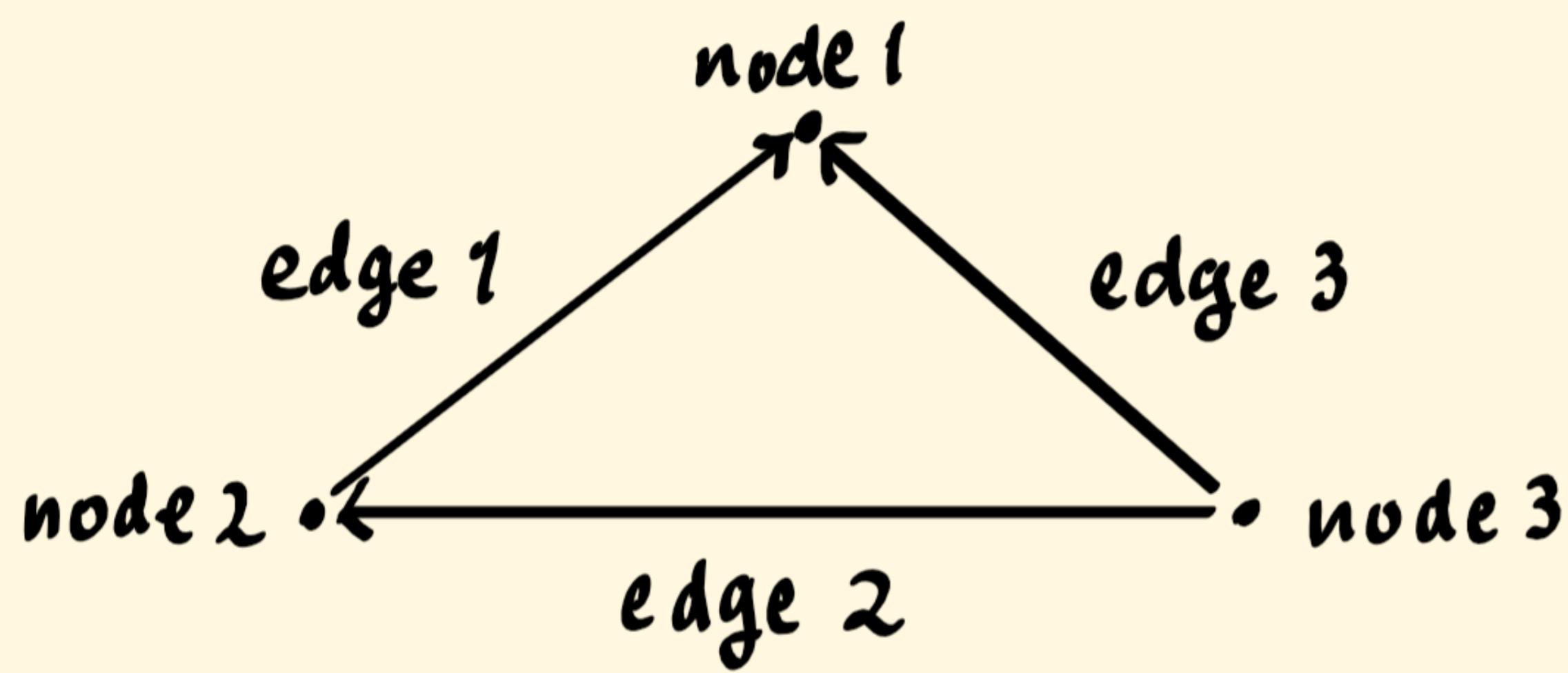
Note that (3) implies we are assuming  $n \leq m$ . If  $n > m$  then  $r=n$  is not possible. We could not, for example, find 5 linearly independent vectors of length 4.

Proof: Suppose the columns  $a_1, \dots, a_n$  of  $A$  are independent.

- (1)  $Ax=0 \rightarrow x_1a_1 + \dots + x_na_n = 0 \rightarrow x_i=0 \ \forall i \rightarrow x=0$ .
- (2)  $Ax=b = Ay \rightarrow A(x-y) = b - b = 0 \rightarrow x-y=0$  by (1)  $\rightarrow x=y$
- (3)  $A$  has  $n$  columns. We assumed these columns are independent.  
The rank is defined to be the number of independent columns.  $\therefore r=n$ .

## Incidence Matrices

Graph:



Edge-node incidence matrix:

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

rows correspond to edges  
columns correspond to nodes  
+1 for an edge entering a node  
-1 for an edge leaving a node

$Ax = b$  :

$$\left. \begin{array}{l} x_1 - x_2 = b_1 \\ x_2 - x_3 = b_2 \\ x_1 - x_3 = b_3 \end{array} \right\} \rightarrow (x_1 - x_2) + (x_2 - x_3) = (x_1 - x_3) \\ b_1 + b_2 = b_3$$

For  $Ax = b$  to have a solution, we must have  $y^T b = 0$ ,  $y = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

This is an example of Kirchhoff's voltage law: the sum of the potential drops around a closed loop is zero.

Consider  $A^T y = f$ . Call  $A^T$  the node-edge incidence matrix.

$$A^T y = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} y_1 + y_3 \\ -y_1 + y_2 \\ -y_2 - y_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = f$$

Questions about  $A^T y = f$ :

(1) For which  $f$  can the system be solved?

$$(y_1 + y_3) + (-y_1 + y_2) - (y_2 + y_3) = 0 \rightarrow f_1 + f_2 + f_3 = 0$$

(The  $y$  components sum to 0  $\rightarrow$  The  $f$  components sum to 0)

The column space of  $A^T$ ,  $R(A^T)$  contains vectors  $f$  s.t.  
 $x^T f = 0$ ,  $x^T = [1 \ 1 \ 1]$ .

(2) Are the solutions to  $A^T y = f$  unique?

Any vector  $y_0 = [c, c, -c]^T$  gives  $A^T y_0 = 0$  and so for any solution  $y$  of  $A^T y = f$ ,  $A^T(y + y_0) = f$  as well. The solution of  $A^T y$  is not unique.

Note that  $x = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$  is in  $N(A)$  and is perpendicular to the columns of  $A^T$ .

Note that  $y = \begin{bmatrix} 1 \\ -1 \\ \vdots \\ 0 \end{bmatrix}$  is in  $N(A^T)$  and is perpendicular to the columns of  $A$ .

From each matrix  $A$  come four fundamental subspaces, two from  $Ax = b$  and two from  $A^T y = f$ :

$R(A)$ : column space of  $A$ , a subspace of  $\mathbb{R}^m$

$N(A)$ : nullspace of  $A$ , a subspace of  $\mathbb{R}^n$

$R(A^T)$ : row space of  $A$ , a subspace of  $\mathbb{R}^n$

$N(A^T)$ : left nullspace of  $A$ , a subspace of  $\mathbb{R}^m$

The dimensions and perpendicularity of these four spaces make up:

### 1P (The Fundamental Theorem of Linear Algebra)

The row space and null space are perpendicular to each other and their dimensions add up to  $n$ :

$$(i) R(A^T) \perp N(A)$$

$$(ii) \dim R(A^T) + \dim N(A^T) = r + (n-r) = n$$

The column space and left nullspace are perpendicular to each other and their dimensions add up to  $m$ :

$$(iii) R(A) \perp N(A^T)$$

$$(iv) \dim R(A) + \dim N(A^T) = r + (m-r) = m$$

Proof:

$$\left. \begin{array}{l} (i) f \in R(A^T) \rightarrow \exists y \text{ s.t. } A^T y = f \\ x \in N(A) \rightarrow A x = 0 \end{array} \right\} \rightarrow x^T f = x^T A^T y = (Ax)^T y = 0^T y = 0$$

$$\left. \begin{array}{l} (iii) b \in R(A) \rightarrow \exists x \text{ s.t. } A x = b \\ y \in N(A^T) \rightarrow A^T y = 0 \end{array} \right\} \rightarrow y^T b = y^T A x = (A^T y)^T x = 0^T x = 0$$

The theorem claims  $\dim R(A) = \dim R(A^T) = r$ . That is, the number of independent columns of  $A$  equals the number of independent rows of  $A$ . One proof of this fact uses orthogonality (orthogonal = perpendicular here).

Let  $A \in \mathbb{R}^{m \times n}$  with  $\dim R(A^T) = r$ . Let  $x_1, \dots, x_r$  be a basis for  $R(A^T)$ .

$$\text{Consider } 0 = c_1 A x_1 + \dots + c_r A x_r = A(c_1 x_1 + \dots + c_r x_r) = A v$$

$$\left. \begin{array}{l} \text{(a) } v = c_1 x_1 + \dots + c_r x_r \text{ means } v \in R(A^T) \\ \text{(b) } A v = 0 \text{ means } v \text{ is orthogonal to each row of } A \\ \text{and thus orthogonal to any linear combination of the} \\ \text{rows of } A (= R(A)) \end{array} \right\} \begin{array}{l} v \in R(A^T) \\ \$ \\ v \perp R(A^T) \\ \rightarrow v = 0 \end{array}$$

$$\therefore c_1 x_1 + \dots + c_r x_r = 0 \rightarrow c_1 = \dots = c_r = 0 \text{ since } x_1, \dots, x_r \text{ are independent.}$$

This means  $A x_1, \dots, A x_r$  are independent as well. Each  $A x_i$  is a vector in  $R(A)$ , so we have found  $r$  independent vectors in  $R(A)$ .

$$\dim R(A) \geq r = \dim R(A^T)$$

Next suppose  $\dim R(A) = s$  and let  $x_1, \dots, x_s$  be a basis for  $R(A)$ .

Consider  $0 = c_1 A^T x_1 + \dots + c_s A^T x_s = A^T(c_1 x_1 + \dots + c_s x_s) = A^T u$ . We have

$$\left. \begin{array}{l} \text{(a) } u \in R(A) \\ \text{(b) } u \perp R(A) \end{array} \right\} \rightarrow u = 0 \rightarrow c_1 = \dots = c_s = 0 \rightarrow A^T x_1, \dots, A^T x_s \text{ independent}$$

Since we have  $s$  independent vectors in  $R(A^T)$ ,

$$\dim R(A) = s \leq \dim R(A^T)$$

Conclude that since  $\dim R(A) \geq \dim R(A^T)$  and  $\dim R(A) \leq \dim R(A^T)$ , we have  $s = r$  and

$$\dim R(A) = \dim R(A^T) = r$$

Finally

(ii) If  $\dim R(A^T) = r$ ,  $A^T$  has  $r$  independent columns. Then the nullspace of  $A$  has a basis with  $n-r$  vectors and  $r + (n-r) = n$

(iv) If  $\dim R(A) = r$ ,  $A$  has  $r$  independent columns. Then the nullspace of  $A^T$  has a basis with  $m-r$  vectors and  $r + (m-r) = m$ .

Another way to see the orthogonality relationships  $R(A^T) \perp N(A)$  and  $R(A) \perp N(A^T)$  is:

$$x \in N(A) \iff Ax = 0$$

$$\iff \begin{bmatrix} \text{row 1} \\ \vdots \\ \text{row } m \end{bmatrix} x = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\iff x$  orthogonal to each row

$\iff x$  orthogonal to any linear combination of rows

$\iff x$  orthogonal to each  $f \in R(A^T)$

$$y \in N(A^T) \iff A^T y = 0$$

$$\iff \begin{bmatrix} \text{column 1} \\ \vdots \\ \text{column } n \end{bmatrix} y = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\iff y$  orthogonal to each column

$\iff y$  orthogonal to any linear combination of columns

$\iff y$  orthogonal to each  $b \in R(A)$

## Exercises

### 1.6.19 True or False?

1. There is no matrix  $A$  whose row space contains  $[1 \ 2 \ 1 \ 1]^T$  and whose nullspace contains  $[1 \ -2 \ 1 \ 1]^T$ .

True

Suppose  $f = [1 \ 2 \ 1 \ 1]^T \in R(A^T)$  and  $x = [1 \ -2 \ 1 \ 1]^T \in N(A)$ . By the fundamental theorem of linear algebra,  $R(A^T) \perp N(A)$ . We must have  $x^T f = 0$ . But  $x^T f = -1$ . From this contradiction conclude there can be no  $A$  s.t.  $f \in R(A^T)$  and  $x \in N(A)$ .

2. Exactly one vector is in both the row space and the row space.

True

Suppose  $x \in R(A^T)$  and  $x \in N(A)$ . Since  $x \in N(A)$ ,  $Ax = 0$ . Then  $x$  is orthogonal to every row of  $A$  and therefore orthogonal to every linear combination of rows of  $A$ . This means  $x$  is orthogonal to every vector in  $R(A^T)$ . In particular  $x \perp x$ . The only vector orthogonal to itself in any vector space is the zero vector. There is exactly one zero vector in  $\mathbb{R}^n$  (or any vector space for that matter).

3. If  $\text{rank } A = \text{rank } B = 3$ , then  $\text{rank}(A+B) \leq 6$ .

True

In general,  $\text{rank}(A+B) \leq \text{rank } A + \text{rank } B$ . The rank of  $A+B$  is the dimension of the column space of  $A+B$ . If any vector  $y$  is in the column space of  $A+B$ , then  $y$  can be written as a linear combination of the vectors  $a_1+b_1, a_2+b_2, \dots, a_n+b_n$ . But this means  $y$  can be written as a linear combination of the vectors  $a_1, \dots, a_n, b_1, \dots, b_n$ . This means the column space of  $A+B$  is a subset of the space spanned by  $a_1, \dots, a_n, b_1, \dots, b_n$ . Then  $\text{rank}(A+B)$ , which is the dimension of the space spanned by the columns of  $A+B$ , is less than the dimension of the space spanned by  $a_1, \dots, a_n, b_1, \dots, b_n$  (since the first is a subset of the second). We know there are  $\text{rank } A + \text{rank } B$  independent vectors out of  $a_1, \dots, a_n, b_1, \dots, b_n$  (= dimension of the span of these vectors). So  $\text{rank}(A+B) \leq \text{rank } A + \text{rank } B$ . Conclude

$$\text{rank}(A+B) \leq \text{rank } A + \text{rank } B = 6.$$

4. The rank of the matrix with every  $a_{ij}=1$  is 1.

True

In this case every column is the same - only 1 independent column.

5. The rank of the  $n \times n$  matrix A with  $a_{ij}=i+j$  is  $r=n$ .

False

Consider the counterexample given by the case  $n=3$ :

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}$$

Suppose  $Ax=0$ :  $\begin{cases} 2x_1 + 3x_2 + 4x_3 = 0 \\ 3x_1 + 4x_2 + 5x_3 = 0 \\ 4x_1 + 5x_2 + 6x_3 = 0 \end{cases}$  If  $x_1 + x_2 + x_3 = 0$  and  $2x_1 + 3x_2 + 4x_3 = 0$  then all 3 equations are satisfied.

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 3 & 4 & 0 \end{array} \right] \sim \left[ \begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array} \right] \rightarrow x_1 = x_3 \text{ and } x_2 = -2x_3$$

Take  $x = [1 \ -2 \ 1]^T$ . Then  $Ax=0$  for this nonzero  $x$ . Conclude that  $r=\text{rank } A < 3 = n$ .