Appendix A Proofs

Theorem A.1 (Schwarz inequality) $\forall$ two $n-$dimensional vectors $\mathbf{v}$, $\mathbf{w}$ where $n \in \mathbb{N}^+$:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

There is

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \ \|\mathbf{w}\|.$$

Proof.

1. **Basis**

   For $n = 1$, there is

   $$|\mathbf{v} \cdot \mathbf{w}| = |v_1 w_1| = |v_1| \ |w_1| = \|\mathbf{v}\| \ \|\mathbf{w}\|.$$

   For $n = 2$, there is

   $$|\mathbf{v} \cdot \mathbf{w}| = |v_1 w_1 + v_2 w_2| \geq 0,$$

   and

   $$\|\mathbf{v}\| \ \|\mathbf{w}\| = \sqrt{v_1^2 + v_2^2}\sqrt{w_1^2 + w_2^2} \geq 0.$$

   Then

   $$|\mathbf{v} \cdot \mathbf{w}|^2 = |v_1 w_1 + v_2 w_2|^2 = v_1^2 w_1^2 + v_2^2 w_2^2 + 2 v_1 v_2 w_1 w_2,$$

   and

   $$(\|\mathbf{v}\| \ \|\mathbf{w}\|)^2 = (v_1^2 + v_2^2)(w_1^2 + w_2^2) = v_1^2 w_1^2 + v_2^2 w_2^2 + v_1^2 w_2^2 + v_2^2 w_1^2.$$

   Thus,

   $$0 \leq (v_1 w_2 - v_2 w_1)^2 \quad \Rightarrow \quad 2 v_1 v_2 w_1 w_2 \leq v_1^2 w_2^2 + v_2^2 w_1^2$$

$$\Rightarrow \quad |\mathbf{v}\cdot\mathbf{w}|^2 \le (\|\mathbf{v}\|\ \|\mathbf{w}\|)^2$$
$$\Rightarrow \quad |\mathbf{v}\cdot\mathbf{w}| \le \|\mathbf{v}\|\ \|\mathbf{w}\|.$$

2. **Induction Assumption**

   Assume that for $n = i$ where $i = 1, 2, \cdots$ and $i \in \mathbb{N}^+$, there is

   $$|\mathbf{v}\cdot\mathbf{w}| \le \|\mathbf{v}\|\ \|\mathbf{w}\|.$$

3. **Induction Step**

   Assume that for $n = i+1$ where $i = 1, 2, \cdots$ and $i \in \mathbb{N}^+$, let

   $$\mathbf{v}_0 = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_i \end{bmatrix}, \qquad \mathbf{w}_0 = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_i \end{bmatrix}.$$

   From the induction Assumption, there is

   $$|\mathbf{v}_0\cdot\mathbf{w}_0| \le \|\mathbf{v}_0\|\ \|\mathbf{w}_0\|.$$

   For $\mathbf{v}$ and $\mathbf{w}$, there is

   $$\begin{aligned} |\mathbf{v}\cdot\mathbf{w}| &= |v_1w_1 + v_2w_2 + \cdots + v_iw_i + v_{i+1}w_{i+1}| \\ &\le |v_1w_1 + v_2w_2 + \cdots + v_iw_i| + |v_{i+1}w_{i+1}| \\ &= |\mathbf{v}_0\cdot\mathbf{w}_0| + |v_{i+1}w_{i+1}| \\ &\le \|\mathbf{v}_0\|\ \|\mathbf{w}_0\| + |v_{i+1}w_{i+1}| \\ &= \|\mathbf{v}_0\|\ \|\mathbf{w}_0\| + \sqrt{v_{i+1}^2w_{i+1}^2}. \end{aligned}$$

   Then

   $$\begin{aligned} |\mathbf{v}\cdot\mathbf{w}|^2 &\le \left(\|\mathbf{v}_0\|\ \|\mathbf{w}_0\| + \sqrt{v_{i+1}^2w_{i+1}^2}\right)^2 \\ &= (\|\mathbf{v}_0\|\ \|\mathbf{w}_0\|)^2 + \left(\sqrt{v_{i+1}^2w_{i+1}^2}\right)^2 + 2\|\mathbf{v}_0\|\ \|\mathbf{w}_0\|\ \sqrt{v_{i+1}^2w_{i+1}^2} \\ &\le \|\mathbf{v}_0\|^2\ \|\mathbf{w}_0\|^2 + v_{i+1}^2w_{i+1}^2 + \left(\|\mathbf{v}_0\|\ \sqrt{w_{i+1}^2}\right)^2 + \left(\|\mathbf{w}_0\|\ \sqrt{v_{i+1}^2}\right)^2 \\ &= \|\mathbf{v}_0\|^2\ \|\mathbf{w}_0\|^2 + v_{i+1}^2w_{i+1}^2 + \|\mathbf{v}_0\|^2w_{i+1}^2 + \|\mathbf{w}_0\|^2v_{i+1}^2 \end{aligned}$$

$$
\begin{aligned}
&= (\|\mathbf{v}_0\|^2 + v_{i+1}^2)(\|\mathbf{w}_0\|^2 + w_{i+1}^2) \\
&= (v_1^2 + v_2^2 + \cdots + v_i^2 + v_{i+1}^2)(w_1^2 + w_2^2 + \cdots + w_i^2 + w_{i+1}^2) \\
&= \|\mathbf{v}\|^2 \ \|\mathbf{w}\|^2 \\
&= (\|\mathbf{v}\| \ \|\mathbf{w}\|)^2.
\end{aligned}
$$

Hence, with $|\mathbf{v} \cdot \mathbf{w}| > 0$ and $\|\mathbf{v}\| \ \|\mathbf{w}\| > 0$, there is

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \ \|\mathbf{w}\|.$$

To sum up, $\forall$ two $n-$dimensional vectors $\mathbf{v}$, $\mathbf{w}$ where $n \in \mathbb{N}^+$, there is

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \ \|\mathbf{w}\|.$$

■

Theorem A.2 (Triangle inequality) $\forall$ two $n-$dimensional vectors $\mathbf{v}$, $\mathbf{w}$ where $n \in \mathbb{N}^+$:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

There is

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

Proof.

1. **Basis**

   For $n = 1$, there is

   $$\|\mathbf{v} + \mathbf{w}\| = |v_1 + w_1| \leq |v_1| + |w_1| = \|\mathbf{v}\| + \|\mathbf{w}\|.$$

   For $n = 2$, there is

   $$\|\mathbf{v} + \mathbf{w}\| = \sqrt{(v_1 + w_1)^2 + (v_2 + w_2)^2} \geq 0,$$

   and

   $$\|\mathbf{v}\| + \|\mathbf{w}\| = \sqrt{v_1^2 + v_2^2} + \sqrt{w_1^2 + w_2^2} \geq 0.$$

Then

$$\begin{aligned}
\|\mathbf{v}+\mathbf{w}\|^2 &= (v_1+w_1)^2+(v_2+w_2)^2 \\
&= v_1^2+w_1^2+v_2^2+w_2^2+2v_1w_1+2v_2w_2 \\
&\leq v_1^2+w_1^2+v_2^2+w_2^2+2|v_1w_1+v_2w_2| \\
&= v_1^2+w_1^2+v_2^2+w_2^2+2\sqrt{(v_1w_1+v_2w_2)^2} \\
&= v_1^2+v_2^2+w_1^2+w_2^2+2\sqrt{v_1^2w_1^2+v_2^2w_2^2+2v_1v_2w_1w_2} \\
&\leq v_1^2+v_2^2+w_1^2+w_2^2+2\sqrt{v_1^2w_1^2+v_2^2w_2^2+v_1^2w_2^2+v_2^2w_1^2} \\
&= v_1^2+v_2^2+w_1^2+w_2^2+2\sqrt{v_1^2+v_2^2}\sqrt{w_1^2+w_2^2} \\
&= (\|\mathbf{v}\|+\|\mathbf{w}\|)^2.
\end{aligned}$$

Thus,

$$\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\|+\|\mathbf{w}\|.$$

2. **Induction Assumption**

   Assume that for $n=i$ where $i=1,2,\cdots$ and $i\in\mathbb{N}^+$, there is

   $$\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\|+\|\mathbf{w}\|.$$

3. **Induction Step**

   Assume that for $n=i+1$ where $i=1,2,\cdots$ and $i\in\mathbb{N}^+$, let

   $$\mathbf{v}_0=\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_i \end{bmatrix}, \qquad \mathbf{w}_0=\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_i \end{bmatrix}.$$

   From the induction Assumption, there is

   $$\|\mathbf{v}_0+\mathbf{w}_0\| \leq \|\mathbf{v}_0\|+\|\mathbf{w}_0\|.$$

   For $\mathbf{v}$ and $\mathbf{w}$, there is

   $$\begin{aligned}
   |\mathbf{v}+\mathbf{w}| &= \sqrt{(v_1+w_1)^2+(v_2+w_2)^2+\cdots+(v_i+w_i)^2+(v_{i+1}+w_{i+1})^2} \\
   &= \sqrt{\|\mathbf{v}_0+\mathbf{w}_0\|^2+(v_{i+1}+w_{i+1})^2} \\
   &\leq \sqrt{(\|\mathbf{v}_0\|+\|\mathbf{w}_0\|)^2+(v_{i+1}+w_{i+1})^2}.
   \end{aligned}$$

Then

$$\begin{aligned}
|\mathbf{v}+\mathbf{w}|^2 &\le (\|\mathbf{v}_0\|+\|\mathbf{w}_0\|)^2+(v_{i+1}+w_{i+1})^2\\
&= \|\mathbf{v}_0\|^2+\|\mathbf{w}_0\|^2+2\|\mathbf{v}_0\|\ \|\mathbf{w}_0\|+v_{i+1}^2+w_{i+1}^2+2v_{i+1}w_{i+1}\\
&= (\|\mathbf{v}_0\|^2+v_{i+1}^2)+(\|\mathbf{w}_0\|^2+w_{i+1}^2)+2(\|\mathbf{v}_0\|\ \|\mathbf{w}_0\|+v_{i+1}w_{i+1})\\
&= \|\mathbf{v}\|^2+\|\mathbf{w}\|^2+2\sqrt{(\|\mathbf{v}_0\|\ \|\mathbf{w}_0\|+v_{i+1}w_{i+1})^2}\\
&= \|\mathbf{v}\|^2+\|\mathbf{w}\|^2+2\sqrt{\|\mathbf{v}_0\|^2\|\mathbf{w}_0\|^2+v_{i+1}^2w_{i+1}^2+2\|\mathbf{v}_0\|\ \|\mathbf{w}_0\|v_{i+1}w_{i+1}}\\
&\le \|\mathbf{v}\|^2+\|\mathbf{w}\|^2+2\sqrt{\|\mathbf{v}_0\|^2\|\mathbf{w}_0\|^2+v_{i+1}^2w_{i+1}^2+\|\mathbf{v}_0\|^2w_{i+1}^2+\|\mathbf{w}_0\|^2v_{i+1}^2}\\
&= \|\mathbf{v}\|^2+\|\mathbf{w}\|^2+2\sqrt{\|\mathbf{v}_0\|^2+v_{i+1}^2}\sqrt{\|\mathbf{w}_0\|^2+w_{i+1}^2}\\
&= \|\mathbf{v}\|^2+\|\mathbf{w}\|^2+2\|\mathbf{v}\|\ \|\mathbf{w}\|\\
&= (\|\mathbf{v}\|+\|\mathbf{w}\|)^2
\end{aligned}$$

Hence, with $|\mathbf{v}+\mathbf{w}|>0$ and $\|\mathbf{v}\|+\|\mathbf{w}\|>0$, there is

$$\|\mathbf{v}+\mathbf{w}\|\le\|\mathbf{v}\|+\|\mathbf{w}\|.$$

To sum up, $\forall$ two $n-$dimensional vectors $\mathbf{v}$, $\mathbf{w}$ where $n\in\mathbb{N}^+$, there is

$$|\mathbf{v}+\mathbf{w}|\le\|\mathbf{v}\|+\|\mathbf{w}\|.$$

■

Theorem A.3 $\forall\ p\times q$ matrix A and $\forall\ q\times r$ matrix B where $p,q,r\in\mathbb{N}^+$

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q}\\ a_{21} & a_{22} & \cdots & a_{2q}\\ \vdots & \vdots & \ddots & \vdots\\ a_{p1} & a_{p2} & \cdots & a_{pq}\end{bmatrix}=\begin{bmatrix} — & \mathbf{a}_{\mathbf{r}_1}^{\mathrm{T}} & —\\ — & \mathbf{a}_{\mathbf{r}_2}^{\mathrm{T}} & —\\ \vdots & \vdots & \vdots\\ — & \mathbf{a}_{\mathbf{r}_p}^{\mathrm{T}} & —\end{bmatrix}=\begin{bmatrix} | & | & \cdots & |\\ \mathbf{a}_{\mathbf{c}_1} & \mathbf{a}_{\mathbf{c}_2} & \cdots & \mathbf{a}_{\mathbf{c}_q}\\ | & | & \cdots & |\end{bmatrix},$$

$$B=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r}\\ b_{21} & b_{22} & \cdots & b_{2r}\\ \vdots & \vdots & \ddots & \vdots\\ b_{q1} & b_{q2} & \cdots & b_{qr}\end{bmatrix}=\begin{bmatrix} — & \mathbf{b}_{\mathbf{r}_1}^{\mathrm{T}} & —\\ — & \mathbf{b}_{\mathbf{r}_2}^{\mathrm{T}} & —\\ \vdots & \vdots & \vdots\\ — & \mathbf{b}_{\mathbf{r}_q}^{\mathrm{T}} & —\end{bmatrix}=\begin{bmatrix} | & | & \cdots & |\\ \mathbf{b}_{\mathbf{c}_1} & \mathbf{b}_{\mathbf{c}_2} & \cdots & \mathbf{b}_{\mathbf{c}_r}\\ | & | & \cdots & |\end{bmatrix}.$$

There are three variants of matrix multiplication AB

1. **Matrix A times every column of matrix B**

$$AB=A\begin{bmatrix} | & | & \cdots & |\\ \mathbf{b}_{\mathbf{c}_1} & \mathbf{b}_{\mathbf{c}_2} & \cdots & \mathbf{b}_{\mathbf{c}_r}\\ | & | & \cdots & |\end{bmatrix}=\begin{bmatrix} | & | & \cdots & |\\ A\mathbf{b}_{\mathbf{c}_1} & A\mathbf{b}_{\mathbf{c}_2} & \cdots & A\mathbf{b}_{\mathbf{c}_r}\\ | & | & \cdots & |\end{bmatrix}.$$

2. **Every row of matrix A times matrix B**

$$AB = \begin{bmatrix} — & \mathbf{a}_{\mathbf{r}_1}^{\mathrm{T}} & — \\ — & \mathbf{a}_{\mathbf{r}_2}^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{a}_{\mathbf{r}_p}^{\mathrm{T}} & — \end{bmatrix} B = \begin{bmatrix} — & \mathbf{a}_{\mathbf{r}_1}^{\mathrm{T}} B & — \\ — & \mathbf{a}_{\mathbf{r}_2}^{\mathrm{T}} B & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{a}_{\mathbf{r}_p}^{\mathrm{T}} B & — \end{bmatrix}.$$

3. **The sum of column i of A times row i of B from 1 to q**

$$AB = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{a}_{\mathbf{c}_1} & \mathbf{a}_{\mathbf{c}_2} & \cdots & \mathbf{a}_{\mathbf{c}_q} \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} — & \mathbf{b}_{\mathbf{r}_1}^{\mathrm{T}} & — \\ — & \mathbf{b}_{\mathbf{r}_2}^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{b}_{\mathbf{r}_q}^{\mathrm{T}} & — \end{bmatrix} = \mathbf{a}_{\mathbf{c}_1}\mathbf{b}_{\mathbf{r}_1}^{\mathrm{T}} + \mathbf{a}_{\mathbf{c}_2}\mathbf{b}_{\mathbf{r}_2}^{\mathrm{T}} + \cdots + \mathbf{a}_{\mathbf{c}_q}\mathbf{b}_{\mathbf{r}_q}^{\mathrm{T}}.$$

Proof.

1. $\forall\ 1 \le j \le r$ and $j \in \mathbb{N}^+$, there is

$$\mathbf{b}_{\mathbf{c}_j} = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{qj} \end{bmatrix}.$$

Then

$$A\mathbf{b}_{\mathbf{c}_i} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{qj} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{q} a_{1i}b_{ij} \\ \sum_{i=1}^{q} a_{2i}b_{ij} \\ \vdots \\ \sum_{i=1}^{q} a_{pi}b_{ij} \end{bmatrix}.$$

Hence

$$\begin{bmatrix} | & | & \cdots & | \\ A\mathbf{b}_{\mathbf{c}_1} & A\mathbf{b}_{\mathbf{c}_2} & \cdots & A\mathbf{b}_{\mathbf{c}_r} \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{q} a_{1i}b_{i1} & \sum_{i=1}^{q} a_{1i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{1i}b_{ir} \\ \sum_{i=1}^{q} a_{2i}b_{i1} & \sum_{i=1}^{q} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{2i}b_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} a_{pi}b_{i1} & \sum_{i=1}^{q} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{pi}b_{ir} \end{bmatrix} = AB.$$

■

2. $\forall\ 1 \leq j \leq p$ and $j \in \mathbb{N}^+$, there is

$$\mathbf{a}_{\mathbf{r}_j} = \begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jq} \end{bmatrix}.$$

Then

$$\begin{aligned} \mathbf{a}_{\mathbf{r}_j}^{\mathrm{T}} B &= \begin{bmatrix} a_{j1} & a_{j2} & \cdots & a_{jq} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q1} & b_{q2} & \cdots & b_{qr} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^{q} a_{ji} b_{i1} & \sum_{i=1}^{q} a_{ji} b_{i2} & \cdots & \sum_{i=1}^{q} a_{ji} b_{ir} \end{bmatrix}. \end{aligned}$$

Hence

$$\begin{bmatrix} — & \mathbf{a}_{\mathbf{r}_1}^{\mathrm{T}} B & — \\ — & \mathbf{a}_{\mathbf{r}_2}^{\mathrm{T}} B & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{a}_{\mathbf{r}_p}^{\mathrm{T}} B & — \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{q} a_{1i} b_{i1} & \sum_{i=1}^{q} a_{1i} b_{i2} & \cdots & \sum_{i=1}^{q} a_{1i} b_{ir} \\ \sum_{i=1}^{q} a_{2i} b_{i1} & \sum_{i=1}^{q} a_{2i} b_{i2} & \cdots & \sum_{i=1}^{q} a_{2i} b_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} a_{pi} b_{i1} & \sum_{i=1}^{q} a_{2i} b_{i2} & \cdots & \sum_{i=1}^{q} a_{pi} b_{ir} \end{bmatrix} = AB.$$

■

3. $\forall\ 1 \leq i \leq q$ and $i \in \mathbb{N}^+$, there are

$$\mathbf{a}_{\mathbf{c}_i} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{pi} \end{bmatrix}, \qquad \mathbf{b}_{\mathbf{r}_i} = \begin{bmatrix} b_{i1} \\ b_{i2} \\ \vdots \\ b_{ir} \end{bmatrix}.$$

Then

$$\mathbf{a}_{\mathbf{c}_i} \mathbf{b}_{\mathbf{r}_i}^{\mathrm{T}} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{pi} \end{bmatrix} \begin{bmatrix} b_{i1} & b_{i2} & \cdots & b_{ir} \end{bmatrix} = \begin{bmatrix} a_{1i} b_{i1} & a_{1i} b_{i2} & \cdots & a_{1i} b_{ir} \\ a_{2i} b_{i1} & a_{2i} b_{i2} & \cdots & a_{2i} b_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ a_{pi} b_{i1} & a_{pi} b_{i2} & \cdots & a_{pi} b_{ir} \end{bmatrix}.$$

Hence

$$\mathbf{a}_{\mathbf{c}_1}\mathbf{b}_{\mathbf{r}_1}^{\mathrm{T}} + \mathbf{a}_{\mathbf{c}_2}\mathbf{b}_{\mathbf{r}_2}^{\mathrm{T}} + \cdots + \mathbf{a}_{\mathbf{c}_q}\mathbf{b}_{\mathbf{r}_q}^{\mathrm{T}} = \begin{bmatrix} \sum_{i=1}^{q} a_{1i}b_{i1} & \sum_{i=1}^{q} a_{1i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{1i}b_{ir} \\ \sum_{i=1}^{q} a_{2i}b_{i1} & \sum_{i=1}^{q} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{2i}b_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} a_{pi}b_{i1} & \sum_{i=1}^{q} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{pi}b_{ir} \end{bmatrix} = AB.$$

■

Theorem A.4 (The Laws of Matrix Operations)

1. **Laws of Matrix Addition**

   $\forall$ number $k \in \mathbb{R}$, $\forall$ three $m \times n$ matrices A, B and C where $m, n \in \mathbb{N}^+$

   - **Commutative Law**

     $$A + B = B + A.$$

   - **Distributive Law**

     $$k(A + B) = kA + kB.$$

   - **Associative Law**

     $$A + (B + C) = (A + B) + C.$$

2. **Laws of Matrix Multiplication**

   - **Distributive Law from The Left**

     $\forall$ $p \times q$ matrix A, $\forall$ $q \times r$ matrices B and C where $p, q, r \in \mathbb{N}^+$

     $$A(B + C) = AB + AC.$$

   - **Distributive Law from The Right**

     $\forall$ $p \times q$ matrices A and B, $\forall$ $q \times r$ matrix C where $p, q, r \in \mathbb{N}^+$

     $$(A + B)C = AC + BC.$$

   - **Associative Law**

     $\forall$ $p \times q$ matrix A, $\forall$ $q \times r$ matrix B and $\forall$ $r \times s$ matrix C where $p, q, r, s \in \mathbb{N}^+$

     $$A(BC) = (AB)C.$$

3. **Laws of Matrix Powers**

   $\forall$ number p, q and $\forall$ $n \times n$ matrix A where $p, q, n \in \mathbb{N}^+$

$$A^p = AA \cdots A \ (p \text{ factors}),$$
$$(A^p)(A^q) = A^{p+q},$$
$$(A^p)^q = A^{pq}.$$

Proof.

1. **Laws of Matrix Addition**

   - **Commutative Law**

     $\forall$ two $m \times n$ matrices A and B where $m, n \in \mathbb{N}^+$ such that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}.$$

     Then

$$\begin{aligned} A + B &= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \\ &= \begin{bmatrix} b_{11} + a_{11} & b_{12} + a_{12} & \cdots & b_{1n} + a_{1n} \\ b_{21} + a_{21} & b_{22} + a_{22} & \cdots & b_{2n} + a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} + a_{m1} & b_{m2} + a_{m2} & \cdots & b_{mn} + a_{mn} \end{bmatrix} \\ &= B + A. \end{aligned}$$

■

   - **Distributive Law**

$\forall$ number $k \in \mathbb{R}$, $\forall$ two $m \times n$ matrices A and B where $m, n \in \mathbb{N}^+$ such that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}.$$

Then

$$\begin{aligned} k(A+B) &= k \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \\ &= \begin{bmatrix} k(a_{11}+b_{11}) & k(a_{12}+b_{12}) & \cdots & k(a_{1n}+b_{1n}) \\ k(a_{21}+b_{21}) & k(a_{22}+b_{22}) & \cdots & k(a_{2n}+b_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ k(a_{m1}+b_{m1}) & k(a_{m2}+b_{m2}) & \cdots & k(a_{mn}+b_{mn}) \end{bmatrix} \\ &= \begin{bmatrix} ka_{11}+kb_{11} & ka_{12}+kb_{12} & \cdots & ka_{1n}+kb_{1n} \\ ka_{21}+kb_{21} & ka_{22}+kb_{22} & \cdots & ka_{2n}+kb_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1}+kb_{m1} & ka_{m2}+kb_{m2} & \cdots & ka_{mn}+kb_{mn} \end{bmatrix} \\ &= \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{bmatrix} + \begin{bmatrix} kb_{11} & kb_{12} & \cdots & kb_{1n} \\ kb_{21} & kb_{22} & \cdots & kb_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ kb_{m1} & kb_{m2} & \cdots & kb_{mn} \end{bmatrix} \\ &= k \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + k \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \\ &= kA + kB. \end{aligned}$$

■

- **Associative Law**

$\forall$ three $m \times n$ matrices A, B and C where $m, n \in \mathbb{N}^+$ such that

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

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix},$$

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}.$$

Then

$$\begin{aligned} A + (B + C) &= A + \begin{bmatrix} b_{11} + c_{11} & b_{12} + c_{12} & \cdots & b_{1n} + c_{1n} \\ b_{21} + c_{21} & b_{22} + c_{22} & \cdots & b_{2n} + c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} + c_{m1} & b_{m2} + c_{m2} & \cdots & b_{mn} + c_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} + b_{11} + c_{11} & a_{12} + b_{12} + c_{12} & \cdots & a_{1n} + b_{1n} + c_{1n} \\ a_{21} + b_{21} + c_{21} & a_{22} + b_{22} + c_{22} & \cdots & a_{2n} + b_{2n} + c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} + c_{m1} & a_{m2} + b_{m2} + c_{m2} & \cdots & a_{mn} + b_{mn} + c_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix} + C \\ &= (A + B) + C. \end{aligned}$$

■

2. **Laws of Matrix Multiplication**

   - **Distributive Law from The Left**

$\forall\ p \times q$ matrix A, $\forall\ q \times r$ matrices B and C where $p, q, r \in \mathbb{N}^+$ such that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix},$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q1} & b_{q2} & \cdots & b_{qr} \end{bmatrix}, \qquad C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1r} \\ c_{21} & c_{22} & \cdots & c_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & \cdots & c_{qr} \end{bmatrix}.$$

Then

$$\begin{aligned} A(B+C) =& A \begin{bmatrix} b_{11}+c_{11} & b_{12}+c_{12} & \cdots & b_{1r}+c_{1r} \\ b_{21}+c_{21} & b_{22}+c_{22} & \cdots & b_{2r}+c_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q1}+c_{q1} & b_{q2}+c_{q2} & \cdots & b_{qr}+c_{qr} \end{bmatrix} \\ =& \begin{bmatrix} \sum_{i=1}^{q} a_{1i}(b_{i1}+c_{i1}) & \sum_{i=1}^{q} a_{1i}(b_{i2}+c_{i2}) & \cdots & \sum_{i=1}^{q} a_{1i}(b_{ir}+c_{ir}) \\ \sum_{i=1}^{q} a_{2i}(b_{i1}+c_{i1}) & \sum_{i=1}^{q} a_{2i}(b_{i2}+c_{i2}) & \cdots & \sum_{i=1}^{q} a_{2i}(b_{ir}+c_{ir}) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} a_{pi}(b_{i1}+c_{i1}) & \sum_{i=1}^{q} a_{pi}(b_{i2}+c_{i2}) & \cdots & \sum_{i=1}^{q} a_{pi}(b_{ir}+c_{ir}) \end{bmatrix} \\ =& \begin{bmatrix} \sum_{i=1}^{q} a_{1i}b_{i1} & \sum_{i=1}^{q} a_{1i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{1i}b_{ir} \\ \sum_{i=1}^{q} a_{2i}b_{i1} & \sum_{i=1}^{q} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{2i}b_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} a_{pi}b_{i1} & \sum_{i=1}^{q} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{q} a_{pi}b_{ir} \end{bmatrix} + \\ & \begin{bmatrix} \sum_{i=1}^{q} a_{1i}c_{i1} & \sum_{i=1}^{q} a_{1i}c_{i2} & \cdots & \sum_{i=1}^{q} a_{1i}c_{ir} \\ \sum_{i=1}^{q} a_{2i}c_{i1} & \sum_{i=1}^{q} a_{2i}c_{i2} & \cdots & \sum_{i=1}^{q} a_{2i}c_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} a_{pi}c_{i1} & \sum_{i=1}^{q} a_{2i}c_{i2} & \cdots & \sum_{i=1}^{q} a_{pi}c_{ir} \end{bmatrix} \\ =& AB + AC. \end{aligned}$$

■

- **Distributive Law from The Right**

  $\forall\ p \times q$ matrices A and B, $\forall\ q \times r$ matrix C where $p, q, r \in \mathbb{N}^+$ such that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pq} \end{bmatrix},$$

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1r} \\ c_{21} & c_{22} & \cdots & c_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & \cdots & c_{qr} \end{bmatrix}.$$

  Then

$$\begin{aligned}
(A+B)C &= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1q}+b_{1q} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2q}+b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}+b_{p1} & a_{p2}+b_{p2} & \cdots & a_{pq}+b_{pq} \end{bmatrix} C \\
&= \begin{bmatrix} \sum_{i=1}^{q}(a_{1i}+b_{1i})c_{i1} & \sum_{i=1}^{q}(a_{1i}+b_{1i})c_{i2} & \cdots & \sum_{i=1}^{q}(a_{1i}+b_{1i})c_{ir} \\ \sum_{i=1}^{q}(a_{2i}+b_{2i})c_{i1} & \sum_{i=1}^{q}(a_{2i}+b_{2i})c_{i2} & \cdots & \sum_{i=1}^{q}(a_{2i}+b_{2i})c_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q}(a_{pi}+b_{pi})c_{i1} & \sum_{i=1}^{q}(a_{pi}+b_{pi})c_{i2} & \cdots & \sum_{i=1}^{q}(a_{pi}+b_{pi})c_{ir} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^{q}a_{1i}c_{i1} & \sum_{i=1}^{q}a_{1i}c_{i2} & \cdots & \sum_{i=1}^{q}a_{1i}c_{ir} \\ \sum_{i=1}^{q}a_{2i}c_{i1} & \sum_{i=1}^{q}a_{2i}c_{i2} & \cdots & \sum_{i=1}^{q}a_{2i}c_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q}a_{pi}c_{i1} & \sum_{i=1}^{q}a_{2i}c_{i2} & \cdots & \sum_{i=1}^{q}a_{pi}c_{ir} \end{bmatrix} + \\
&\quad \begin{bmatrix} \sum_{i=1}^{q}b_{1i}c_{i1} & \sum_{i=1}^{q}b_{1i}c_{i2} & \cdots & \sum_{i=1}^{q}b_{1i}c_{ir} \\ \sum_{i=1}^{q}b_{2i}c_{i1} & \sum_{i=1}^{q}b_{2i}c_{i2} & \cdots & \sum_{i=1}^{q}b_{2i}c_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q}b_{pi}c_{i1} & \sum_{i=1}^{q}b_{2i}c_{i2} & \cdots & \sum_{i=1}^{q}b_{pi}c_{ir} \end{bmatrix} \\
&= AC + BC.
\end{aligned}$$

■

3. **Laws of Matrix Powers**

   $\forall$ number p, q and $\forall$ $n \times n$ matrix A where $p, q, n \in \mathbb{N}^+$, there are

$$(A^p)(A^q) = (\underbrace{AA\cdots A}_{p})(\underbrace{AA\cdots A}_{q}) = \underbrace{AA\cdots A}_{p+q} = A^{p+q},$$

$$(A^p)^q = \underbrace{(\underbrace{AA\cdots A}_{p})(\underbrace{AA\cdots A}_{p})\cdots(\underbrace{AA\cdots A}_{p})}_{q} = \underbrace{AA\cdots A}_{pq} = A^{pq}.$$

■

Theorem A.5 (The Properties of Inverse Matrices)

1. $\forall$ $n \times n$ invertible square matrix A where $n \in \mathbb{N}^+$, its inverse A^{-1} is unique.

2. For $n \times n$ square matrices $A_1, A_2, \cdots, A_k$ where $n, k \in \mathbb{N}^+$, if $A_1, A_2, \cdots, A_k$ are separately invertible, then

$$(A_1 A_2 \cdots A_{k-1} A_k)^{-1} = A_k^{-1} A_{k-1}^{-1} \cdots A_2^{-1} A_1^{-1}.$$

Proof.

1. $\forall$ $n \times n$ invertible square matrix A where $n \in \mathbb{N}^+$, suppose that its inverse A^{-1} is not unique. Assume that the matrices both B and C are its inverse matrices. Then

$$BA = I \quad \Rightarrow \quad BAC = IC \quad \Rightarrow \quad B(AC) = C \quad \Rightarrow \quad BI = C \quad \Rightarrow \quad B = C.$$

   This statement contradicts our assumption. Therefore, $\forall$ $n \times n$ invertible square matrix A where $n \in \mathbb{N}^+$, its inverse A^{-1} is unique.

■

2. For $n \times n$ square matrices $A_1, A_2, \cdots, A_k$ where $n, k \in \mathbb{N}^+$, if $A_1, A_2, \cdots, A_k$ are separately invertible, then

   (a) **Basis**

   For $k = 1$, there is

$$A_1 A_1^{-1} = A_1^{-1} A_1 = I.$$

Thus, with the uniqueness of the inverse, A_1^{-1} is the inverse of A_1.

For $k = 2$, there are

$$(A_1A_2)(A_2^{-1}A_1^{-1}) = A_1(A_2A_2^{-1})A_1^{-1} = A_1IA_1^{-1} = A_1A_1^{-1} = I,$$

and

$$(A_2^{-1}A_1^{-1})(A_1A_2) = A_2^{-1}(A_1^{-1}A_1)A_2 = A_2^{-1}IA_2 = A_2^{-1}A_2 = I.$$

Thus, with the uniqueness of the inverse, $A_2^{-1}A_1^{-1}$ is the inverse of A_1A_2.

(b) **Induction Assumption**

Assume that for $k = i$ where $i = 1, 2, \cdots$ and $i \in \mathbb{N}^+$, there is

$$(A_1A_2 \cdots A_{i-1}A_i)^{-1} = A_i^{-1}A_{i-1}^{-1} \cdots A_2^{-1}A_1^{-1}.$$

(c) **Induction Step**

For $k = i + 1$ where $i = 1, 2, \cdots$ and $i \in \mathbb{N}^+$, there is

$$\begin{aligned}(A_1A_2 \cdots A_{i-1}A_iA_{i+1})^{-1} &= A_{i+1}^{-1}(A_1A_2 \cdots A_{i-1}A_i)^{-1} \\ &= A_{i+1}^{-1}(A_i^{-1}A_{i-1}^{-1} \cdots A_2^{-1}A_1^{-1}) \\ &= A_{i+1}^{-1}A_i^{-1}A_{i-1}^{-1} \cdots A_2^{-1}A_1^{-1}.\end{aligned}$$

To sum up, for $n \times n$ square matrix $A_1, A_2, \cdots, A_k$ where $n, k \in \mathbb{N}^+$, if $A_1, A_2, \cdots, A_k$ are separately invertible, then

$$(A_1A_2 \cdots A_{k-1}A_k)^{-1} = A_k^{-1}A_{k-1}^{-1} \cdots A_2^{-1}A_1^{-1}.$$

■

Theorem A.6 (Calculating Inverse Matrices by Gauss-Jordan Elimination) $\forall\, n \times n$ invertible square matrix A where $n \in \mathbb{N}^+$, apply Gauss-Jordan Elimination into the matrix

$$\begin{bmatrix} A & I \end{bmatrix}$$

to get the matrix

$$\begin{bmatrix} I & A^{-1} \end{bmatrix}.$$

Proof. If $A = I$, then $A^{-1} = I$ and Gauss-Jordan Elimination does nothing. If $A \neq I$, then apply Gauss-Jordan Elimination into the matrix A. Because the RREF of A is I, let

$$R_1 R_2 \cdots R_k A = I.$$

where $R_1, R_2, \cdots, R_k$ are the row operations of Gauss-Jordan Elimination in A and $k \in \mathbb{N}^+$. Then

$$R_1 R_2 \cdots R_k A A^{-1} = I A^{-1} \quad \Rightarrow \quad R_1 R_2 \cdots R_k I = A^{-1}.$$

■

Theorem A.7 (The Properties of Transposes)

1. For $m \times n$ matrices $A_1, A_2, \cdots, A_k$ where $m, n, k \in \mathbb{N}^+$

$$(A_1 + A_2 + \cdots + A_k)^{\mathrm{T}} = A_1^{\mathrm{T}} + A_2^{\mathrm{T}} + \cdots + A_k^{\mathrm{T}}.$$

2. $\forall\ n_1 \times n_2$ matrix A_1, $\forall\ n_2 \times n_3$ matrix A_2, $\cdots$, $\forall\ n_{k-1} \times n_k$ matrix A_{k-1}, $\forall\ n_k \times n_{k+1}$ matrix A_k where $k, n_1, n_2, \cdots, n_k, n_{k+1} \in \mathbb{N}^+$

$$(A_1 A_2 \cdots A_{k-1} A_k)^{\mathrm{T}} = A_k^{\mathrm{T}} A_{k-1}^{\mathrm{T}} \cdots A_2^{\mathrm{T}} A_1^{\mathrm{T}}.$$

3. $\forall\ n \times n$ invertible matrix A where $n \in \mathbb{N}^+$

$$(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1}.$$

Proof.

1. For $1 \leq i \leq k$ where $k \in \mathbb{N}^+$, let

$$A_i = \begin{bmatrix} a_{i11} & a_{i12} & \cdots & a_{i1n} \\ a_{i21} & a_{i22} & \cdots & a_{i2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{im1} & a_{im2} & \cdots & a_{imn} \end{bmatrix}, \qquad A_i^{\mathrm{T}} = \begin{bmatrix} a_{i11} & a_{i21} & \cdots & a_{im1} \\ a_{i12} & a_{i22} & \cdots & a_{im2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1n} & a_{i2n} & \cdots & a_{imn} \end{bmatrix}$$

   Then

$$A_1 + A_2 + \cdots + A_k = \begin{bmatrix} \sum_{i=1}^{k} a_{i11} & \sum_{i=1}^{k} a_{i12} & \cdots & \sum_{i=1}^{k} a_{i1n} \\ \sum_{i=1}^{k} a_{i21} & \sum_{i=1}^{k} a_{i22} & \cdots & \sum_{i=1}^{k} a_{i2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{k} a_{im1} & \sum_{i=1}^{k} a_{im2} & \cdots & \sum_{i=1}^{k} a_{imn} \end{bmatrix}.$$

Hence

$$(A_1 + A_2 + \cdots + A_k)^{\mathrm{T}} = \begin{bmatrix} \sum_{i=1}^{k} a_{i11} & \sum_{i=1}^{k} a_{i21} & \cdots & \sum_{i=1}^{k} a_{im1} \\ \sum_{i=1}^{k} a_{i12} & \sum_{i=1}^{k} a_{i22} & \cdots & \sum_{i=1}^{k} a_{im2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{k} a_{i1n} & \sum_{i=1}^{k} a_{i2n} & \cdots & \sum_{i=1}^{k} a_{imn}. \end{bmatrix}$$
$$= A_1^{\mathrm{T}} + A_2^{\mathrm{T}} + \cdots + A_k^{\mathrm{T}}.$$

■

2. (a) **Basis**

For $k = 1$, $\forall$ $n_1 \times n_2$ matrix A_1 where $n_1, n_2 \in \mathbb{N}^+$, there is

$$A_1^{\mathrm{T}} = A_1^{\mathrm{T}}.$$

For $k = 2$, $\forall$ $n_1 \times n_2$ matrix A_1 and $\forall$ $n_2 \times n_3$ matrix A_2 where $n_1, n_2, n_3 \in \mathbb{N}^+$ such that

$$A_1 = \begin{bmatrix} a_{111} & a_{112} & \cdots & a_{11n_2} \\ a_{121} & a_{122} & \cdots & a_{12n_2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n_11} & a_{1n_12} & \cdots & a_{1n_1n_2} \end{bmatrix}, \qquad A_2 = \begin{bmatrix} a_{211} & a_{212} & \cdots & a_{21n_3} \\ a_{221} & a_{222} & \cdots & a_{22n_3} \\ \vdots & \vdots & \ddots & \vdots \\ a_{2n_21} & a_{2n_22} & \cdots & a_{2n_2n_3} \end{bmatrix}.$$

Then

$$A_1A_2 = \begin{bmatrix} \sum_{i=1}^{n_2} a_{11i}a_{2i1} & \sum_{i=1}^{n_2} a_{11i}a_{2i2} & \cdots & \sum_{i=1}^{n_2} a_{11i}a_{2in_3} \\ \sum_{i=1}^{n_2} a_{12i}a_{2i1} & \sum_{i=1}^{n_2} a_{12i}a_{2i2} & \cdots & \sum_{i=1}^{n_2} a_{12i}a_{2in_3} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n_2} a_{1n_1i}a_{2i1} & \sum_{i=1}^{n_2} a_{1n_1i}a_{2i2} & \cdots & \sum_{i=1}^{n_2} a_{1n_1i}a_{2in_3} \end{bmatrix}$$

and

$$(A_1A_2)^{\mathrm{T}} = \begin{bmatrix} \sum_{i=1}^{n_2} a_{11i}a_{2i1} & \sum_{i=1}^{n_2} a_{12i}a_{2i2} & \cdots & \sum_{i=1}^{n_2} a_{1n_1i}a_{2i1} \\ \sum_{i=1}^{n_2} a_{11i}a_{2i2} & \sum_{i=1}^{n_2} a_{12i}a_{2i2} & \cdots & \sum_{i=1}^{n_2} a_{1n_1i}a_{2i2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n_2} a_{11i}a_{2in_3} & \sum_{i=1}^{n_2} a_{12i}a_{2in_3} & \cdots & \sum_{i=1}^{n_2} a_{1n_1i}a_{2in_3} \end{bmatrix}.$$

Moreover

$$A_1^{\mathrm{T}} = \begin{bmatrix} a_{111} & a_{121} & \cdots & a_{1n_11} \\ a_{112} & a_{122} & \cdots & a_{1n_12} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11n_2} & a_{12n_2} & \cdots & a_{1n_1n_2} \end{bmatrix}, \qquad A_2^{\mathrm{T}} = \begin{bmatrix} a_{211} & a_{221} & \cdots & a_{2n_21} \\ a_{212} & a_{222} & \cdots & a_{2n_22} \\ \vdots & \vdots & \ddots & \vdots \\ a_{21n_3} & a_{22n_3} & \cdots & a_{2n_2n_3} \end{bmatrix},$$

and

$$\begin{aligned} A_2^{\mathrm{T}} A_1^{\mathrm{T}} &= \begin{bmatrix} \sum_{i=1}^{n_2} a_{2i1}a_{11i} & \sum_{i=1}^{n_2} a_{2i2}a_{12i} & \cdots & \sum_{i=1}^{n_2} a_{2i1}a_{1n_1i} \\ \sum_{i=1}^{n_2} a_{2i2}a_{11i} & \sum_{i=1}^{n_2} a_{2i2}a_{12i} & \cdots & \sum_{i=1}^{n_2} a_{2i2}a_{1n_1i} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n_2} a_{2in_3}a_{11i} & \sum_{i=1}^{n_2} a_{2in_3}a_{12i} & \cdots & \sum_{i=1}^{n_2} a_{2in_3}a_{1n_1i} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^{n_2} a_{11i}a_{2i1} & \sum_{i=1}^{n_2} a_{12i}a_{2i2} & \cdots & \sum_{i=1}^{n_2} a_{1n_1i}a_{2i1} \\ \sum_{i=1}^{n_2} a_{11i}a_{2i2} & \sum_{i=1}^{n_2} a_{12i}a_{2i2} & \cdots & \sum_{i=1}^{n_2} a_{1n_1i}a_{2i2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n_2} a_{11i}a_{2in_3} & \sum_{i=1}^{n_2} a_{12i}a_{2in_3} & \cdots & \sum_{i=1}^{n_2} a_{1n_1i}a_{2in_3} \end{bmatrix} \\ &= (A_1A_2)^{\mathrm{T}}. \end{aligned}$$

(b) **Induction Assumption**

Assume that for $k = i$ where $i = 1, 2, \cdots$ and $i \in \mathbb{N}^+$, $\forall$ $n_1 \times n_2$ matrix A_1, $\forall$ $n_2 \times n_3$ matrix A_2, $\cdots$, $\forall$ $n_{i-1} \times n_i$ matrix A_{i-1}, $\forall$ $n_i \times n_{i+1}$ matrix A_i where $i, n_1, n_2, \cdots, n_i, n_{i+1} \in \mathbb{N}^+$, there is

$$(A_1A_2 \cdots A_{i-1}A_i)^{\mathrm{T}} = A_i^{\mathrm{T}} A_{i-1}^{\mathrm{T}} \cdots A_2^{\mathrm{T}} A_1^{\mathrm{T}}.$$

(c) **Induction Step**

For $k = i + 1$ where $i = 1, 2, \cdots$ and $i \in \mathbb{N}^+$, $\forall$ $n_1 \times n_2$ matrix A_1, $\forall$ $n_2 \times n_3$ matrix A_2, $\cdots$, $\forall$ $n_i \times n_{i+1}$ matrix A_i, $\forall$ $n_{i+1} \times n_{i+2}$ matrix A_{i+1} where $i, n_1, n_2, \cdots, n_{i+1}, n_{i+2} \in \mathbb{N}^+$, there is

$$\begin{aligned} (A_1A_2 \cdots A_{i-1}A_iA_{i+1})^{\mathrm{T}} &= A_{i+1}^{\mathrm{T}}(A_iA_{i-1} \cdots A_2A_1)^{\mathrm{T}} \\ &= A_{i+1}^{\mathrm{T}}(A_i^{\mathrm{T}} A_{i-1}^{\mathrm{T}} \cdots A_2^{\mathrm{T}} A_1^{\mathrm{T}}) \\ &= A_{i+1}^{\mathrm{T}} A_i^{\mathrm{T}} A_{i-1}^{\mathrm{T}} \cdots A_2^{\mathrm{T}} A_1^{\mathrm{T}}. \end{aligned}$$

To sum up, $\forall$ $n_1 \times n_2$ matrix A_1, $\forall$ $n_2 \times n_3$ matrix A_2, $\cdots$, $\forall$ $n_{k-1} \times n_k$ matrix A_{k-1}, $\forall$ $n_k \times n_{k+1}$ matrix A_k where $k, n_1, n_2, \cdots, n_k, n_{k+1} \in \mathbb{N}^+$

$$(A_1 A_2 \cdots A_{k-1} A_k)^{\mathrm{T}} = A_k^{\mathrm{T}} A_{k-1}^{\mathrm{T}} \cdots A_2^{\mathrm{T}} A_1^{\mathrm{T}}.$$

■

3. (a) **Existence** (?)

   $\forall$ $n \times n$ matrix A where $n \in \mathbb{N}^+$, if the inverse of A exists, then the rank of A is n. Because A and A^{T} have the same rank, the rank of A^{T} is also n. Therefore, the inverse of A^{T} exists.

   (b) **Equivalence**

   $\forall$ $n \times n$ invertible matrix A where $n \in \mathbb{N}^+$, there is

$$(A^{-1})^{\mathrm{T}} A^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = I^{\mathrm{T}} = I \quad \Rightarrow \quad (A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1}.$$

■

Theorem A.8 The inverse of a symmetric matrix is also symmetric.

Proof. $\forall$ $n \times n$ invertible symmetric matrix S where $n \in \mathbb{N}^+$, there is $S^{\mathrm{T}} = S$. Then

$$(S^{-1})^{\mathrm{T}} = (S^{\mathrm{T}})^{-1} = S^{-1}.$$

Therefore, S^{-1} is also symmetric.

■

Theorem A.9 (Properties of Vector Space and Subspace)

1. Every space and subspace contains the zero vector.
2. Lines through the origin are also subspaces.
3. **Closure Property**

   If $\forall$ vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are in a vector space or subspace and $\forall$ $\alpha_1, \alpha_2 \in \mathbb{R}$, then the linear combination $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$ must stay in this vector space or subspace.

Proof.

1. $\forall$ space $\mathbf{S}$, if every vector in it all have n components, then these components can all be 0. Therefore, $\mathbf{0} \in \mathbf{S}$.

   $\forall$ subspace $\mathbf{S}$, if two vectors $\mathbf{v}, \mathbf{w} \in \mathbf{S}$, then $\mathbf{0} = 0\mathbf{v} + 0\mathbf{w} \in \mathbf{S}$.

   ■

2. $\forall$ line through the origin, there must be some vectors in it. If two vectors $\mathbf{v}$ and $\mathbf{w}$ are in this line, we have that both $\mathbf{v} + \mathbf{w}$ and $k\mathbf{v}$ are in this line where $k \in \mathbb{R}$. Therefore, this line is a subspace.

   ■

3. $\forall$ space or subspace $\mathbf{S}$, if two vectors $\mathbf{v}, \mathbf{w} \in \mathbf{S}$, then $\forall\ \alpha_1, \alpha_2 \in \mathbb{R}$, we have $\alpha_1\mathbf{v} \in \mathbf{S}$ and $\alpha_2\mathbf{w} \in \mathbf{S}$, so $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 \in \mathbf{S}$.

   ■

Theorem A.10 (Properties of Linear Independence)

1. **Zero Vector**

   (1) For vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ where $n \in \mathbb{N}^+$, if $\mathbf{v}_i = \mathbf{0}$ where $1 \leq i \leq n$ and $i \in \mathbb{N}^+$, then $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

   (2) For vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ where $n \in \mathbb{N}^+$, if $\mathbf{v}_i = \mathbf{v}_j$ where $1 \leq i < j \leq n$ and $i, j \in \mathbb{N}^+$, then $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

2. **Contraction and Extension**

   (1)
   - $\forall$ linearly dependent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ and $\forall$ vectors $\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \cdots, \mathbf{v}_m$ where $m > n$ and $m, n \in \mathbb{N}^+$, the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n, \mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \cdots, \mathbf{v}_m$ are linearly dependent.
   - $\forall$ linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ where $n \in \mathbb{N}^+$, the vectors consisting of a subset are also linearly independent.

   (2)
   - $\forall$ linearly dependent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n \in \mathbf{R}^m$ where $m, n \in \mathbb{N}^+$, if every vector adds a component at the same place and they become a series of new vectors $\mathbf{v}_1', \mathbf{v}_2', \cdots, \mathbf{v}_n' \in \mathbf{R}^{m+1}$, then the new vectors are also linearly dependent.

- $\forall$ linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n \in \mathbf{R}^m$ where $m > 1$ and $m, n \in \mathbb{N}^+$, if every vector deletes a component at the same place and they become a series of new vectors $\mathbf{v}'_1, \mathbf{v}'_2, \cdots, \mathbf{v}'_n \in \mathbf{R}^{m-1}$, then the new vectors are also linearly independent.

(3) $\forall$ linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$, if there exists a vector $\mathbf{u}$ such that $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent, then $\mathbf{u}$ can be represented as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ and the scalars are unique.

3. **Core Property of Linear Independence**

For two sets of vectors $S_1 = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$ and $S_2 = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$, where $m, n \in \mathbb{N}^+$, and both S_1 and S_2 are linearly independent.

(1) If every vector in S_2 can be represented as a linear combination of the vectors in S_1, then $m \geq n$.

(2) If every vector in S_2 can be represented as a linear combination of the vectors in S_1 and every vector in S_1 can be represented as a linear combination of the vectors in S_2, then $m = n$.

4. **Maximal Linearly Independent Subset**

(1) There must exist a maximal linearly independent subset in every set of vectors.

(2) For a set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ where $n \in \mathbb{N}^+$, assume that a subset $S' = \{\mathbf{v}_{k_1}, \mathbf{v}_{k_2}, \cdots, \mathbf{v}_{k_m}\}$ of S is a maximal linearly independent subset of S, where $1 \leq m \leq n, 1 \leq k_1 \leq k_2 \leq \cdots \leq k_m \leq n$ and $m, k_1, k_2, \cdots, k_m \in \mathbb{N}^+$. $\forall$ vector $\mathbf{v} \in S$, then $\mathbf{v}, \mathbf{v}_{k_1}, \mathbf{v}_{k_2}, \cdots, \mathbf{v}_{k_m}$ are linearly dependent.

(3) $\forall$ set of vectors, every maximal linearly independent subset of it has the same number of vectors.

5. **Number Restriction and Rank**

For two sets of vectors $S_1 = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$ and $S_2 = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ such that every vector in S_2 can be represented as a linear combination of the vectors in S_1, where $m, n \in \mathbb{N}^+$.

- If $m < n$, then the vectors in S_2 must be linearly dependent.

- If the vectors in S_2 are linearly independent, then $m \geq n$.
- Assume that the rank of S_1 is r_1 and the rank of S_2 is r_2, then $r_1 \geq r_2$.

6. **Basis and Dimension**

   (1) Assume that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ is a basis for a vector space,

   - $\forall$ vector $\mathbf{u}$ in this vector space can be represented as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ and the scalars are unique.
   - $\forall$ vector $\mathbf{u}$ in this vector space such that $\mathbf{u} \neq \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$, then $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

   (2)
   - $\forall\ m, n \in \mathbb{N}^+$, if both $S_1 = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$ and $S_2 = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ are bases for the same vector space, then $m = n$.
   - $\forall\ n \in \mathbb{N}^+$, $\forall\ n$ independent vectors in $\mathbf{R}^n$ must span $\mathbf{R}^n$, which means that they are a basis.

   (3)
   - $\forall\ n \in \mathbb{N}^+$, the dimension of $\mathbf{R}^n$ is n.
   - $\forall\ m, n \in \mathbb{N}^+$ such that $m < n$, $\forall$ vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n \in \mathbf{R}^m$ must be linearly dependent.

   (4) The row operations do not change the dimension of the row space of matrices.

Proof.

1. **Zero vector**

   (1) There is a combination that gives zero vector is

   $$0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_n = \mathbf{0}.$$

   Therefore, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

   ■

   (2) There is a combination that gives zero vector is

   $$0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_{j-1} - 1\mathbf{v}_j + 0\mathbf{v}_{j+1} + \cdots + 0\mathbf{v}_n = \mathbf{0}.$$

   Therefore, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

   ■

2. **Contraction and Extension**

(1)

- For the linearly dependent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$, there is a linear combination such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = 0,$$

and

$$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 \neq 0,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$. For the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n, \mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \cdots, \mathbf{v}_m$, there is a combination such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n + 0\mathbf{v}_{n+1} + 0\mathbf{v}_{n+2} + \cdots + 0\mathbf{v}_m = \mathbf{0}.$$

Therefore, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n, \mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \cdots, \mathbf{v}_m$ are linearly dependent.

■

- Assume that there is a subset vectors $\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \cdots, \mathbf{v}_{i_k}$ are linearly dependent where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ and $k, i_1, i_2, \cdots, i_k \in \mathbb{N}^+$. Hence, there is a linear combination such that

$$\alpha_1\mathbf{v}_{i_1} + \alpha_2\mathbf{v}_{i_2} + \cdots + \alpha_k\mathbf{v}_{i_k} = 0,$$

and

$$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_k^2 \neq 0,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_k \in \mathbb{R}$. Let other vectors be $\mathbf{v}_{j_1}, \mathbf{v}_{j_2}, \cdots, \mathbf{v}_{j_{n-k}}$. Then there is a combination such that

$$\alpha_1\mathbf{v}_{i_1} + \alpha_2\mathbf{v}_{i_2} + \cdots + \alpha_k\mathbf{v}_{i_k} + 0\mathbf{v}_{j_1} + 0\mathbf{v}_{j_2} + \cdots + 0\mathbf{v}_{j_{n-k}} = \mathbf{0}.$$

Therefore, the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent, which contradicts the assumption. Thus, the vectors consisting of a subset are also linearly independent.

■

(2) • Assume that

$$\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ \vdots \\ v_{i1} \\ v_{(i+1)1} \\ \vdots \\ v_{m1} \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} v_{12} \\ \vdots \\ v_{i2} \\ v_{(i+1)2} \\ \vdots \\ v_{m2} \end{bmatrix}, \qquad \cdots, \qquad \mathbf{v}_n = \begin{bmatrix} v_{1n} \\ \vdots \\ v_{in} \\ v_{(i+1)n} \\ \vdots \\ v_{mn} \end{bmatrix},$$

where $1 \leq i \leq m$ and $i \in \mathbb{N}^+$. Then

$$\mathbf{v}_1' = \begin{bmatrix} v_{11} \\ \vdots \\ v_{i1} \\ u_1 \\ v_{(i+1)1} \\ \vdots \\ v_{m1} \end{bmatrix}, \qquad \mathbf{v}_2' = \begin{bmatrix} v_{12} \\ \vdots \\ v_{i2} \\ u_2 \\ v_{(i+1)2} \\ \vdots \\ v_{m2} \end{bmatrix}, \qquad \cdots, \qquad \mathbf{v}_n' = \begin{bmatrix} v_{1n} \\ \vdots \\ v_{in} \\ u_n \\ v_{(i+1)n} \\ \vdots \\ v_{mn} \end{bmatrix}.$$

Because the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent, there at least exists a combination such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = 0,$$

and

$$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 \neq 0,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$. Therefore, $\exists\ 1 \leq j \leq m$ and $j \in \mathbb{N}^+$ such that

$$\alpha_1 v_{j1} + \alpha_2 v_{j2} + \cdots + \alpha_n v_{jn} = 0.$$

This equation still works in the new vectors $\mathbf{v}_1', \mathbf{v}_2', \cdots, \mathbf{v}_n'$, so they are also linearly dependent.

■

- Assume that

$$\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ \vdots \\ v_{(i-1)1} \\ v_{i1} \\ v_{(i+1)1} \\ \vdots \\ v_{m1} \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} v_{12} \\ \vdots \\ v_{(i-1)2} \\ v_{i2} \\ v_{(i+1)2} \\ \vdots \\ v_{m2} \end{bmatrix}, \qquad \cdots, \qquad \mathbf{v}_n = \begin{bmatrix} v_{1n} \\ \vdots \\ v_{(i-1)n} \\ v_{in} \\ v_{(i+1)n} \\ \vdots \\ v_{mn} \end{bmatrix},$$

where $1 \leq i \leq m$ and $i \in \mathbb{N}^+$. Then

$$\mathbf{v}_1' = \begin{bmatrix} v_{11} \\ \vdots \\ v_{(i-1)1} \\ v_{(i+1)1} \\ \vdots \\ v_{m1} \end{bmatrix}, \qquad \mathbf{v}_2' = \begin{bmatrix} v_{12} \\ \vdots \\ v_{(i-1)2} \\ v_{(i+1)2} \\ \vdots \\ v_{m2} \end{bmatrix}, \qquad \cdots, \qquad \mathbf{v}_n' = \begin{bmatrix} v_{1n} \\ \vdots \\ v_{(i-1)n} \\ v_{(i+1)n} \\ \vdots \\ v_{mn} \end{bmatrix}.$$

Because the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent, $\forall\, \alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$ such that

$$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 \neq 0,$$

there is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n \neq 0,$$

Therefore, $\forall\ 1 \leq j \leq m$ and $j \in \mathbb{N}^+$ such that

$$\alpha_1 v_{j1} + \alpha_2 v_{j2} + \cdots + \alpha_n v_{jn} \neq 0.$$

This equation still works in the new vectors $\mathbf{v}_1', \mathbf{v}_2', \cdots, \mathbf{v}_n'$, so they are also linearly independent.

■

(3) • **Existence**

Because $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent, there exists a linear combination such that

$$\beta\mathbf{u} + \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = 0,$$

and

$$\beta^2 + \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 \neq 0,$$

where $\beta, \alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$. Assume that $\beta = 0$, we can find a linear combination such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = 0,$$

and

$$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 \neq 0,$$

so the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent, which contradicts the premise. Therefore, $\beta \neq 0$, and there is

$$\mathbf{u} = -\left(\frac{\alpha_1}{\beta}\mathbf{v}_1 + \frac{\alpha_2}{\beta}\mathbf{v}_2 + \cdots + \frac{\alpha_n}{\beta}\mathbf{v}_n\right).$$

- **Uniqueness**

  Assume that there exists

  $$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{u},$$

  and

  $$\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_n\mathbf{v}_n = \mathbf{u},$$

  where $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_n \in \mathbb{R}$. Then there is

  $$(\alpha_1 - \beta_1)\mathbf{v}_1 + (\alpha_2 - \beta_2)\mathbf{v}_2 + \cdots + (\alpha_n - \beta_n)\mathbf{v}_n = \mathbf{0}.$$

  Because the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent, the only solution of this equation is

  $$\alpha_1 - \beta_1 = 0, \quad \alpha_2 - \beta_2 = 0, \quad \cdots, \quad \alpha_n - \beta_n = 0.$$

  Hence

  $$\alpha_1 = \beta_1, \quad \alpha_2 = \beta_2, \quad \cdots, \quad \alpha_n = \beta_n.$$

■

3. **Core Property of Linear Independence**

Let

$$U = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \\ | & | & \cdots & | \end{bmatrix}, \qquad V = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & \cdots & | \end{bmatrix}.$$

$\forall$ vector $\mathbf{a}$ and $\mathbf{b}$

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

It is clear that

$$U\mathbf{a} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_m\mathbf{u}_m,$$
$$V\mathbf{b} = b_1\mathbf{v}_1 + b_2\mathbf{b}_2 + \cdots + b_n\mathbf{v}_n.$$

Because $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$ are linearly independent and $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent, if $\forall$ vector $\mathbf{x}_u$ or $\mathbf{x}_v$ such that

$$U\mathbf{x}_u = \mathbf{0}, \qquad V\mathbf{x}_v = \mathbf{0}.$$

then

$$\mathbf{x}_u = \mathbf{0}, \qquad \mathbf{x}_v = \mathbf{0}.$$

(1) Suppose that $m < n$.

Because $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$ is a basis for the vector space, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ can be expressed as linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$. Assume that

$$\mathbf{v}_1 = a_{11}\mathbf{u}_1 + a_{21}\mathbf{u}_2 + \cdots + a_{m1}\mathbf{u}_m,$$
$$\mathbf{v}_2 = a_{12}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + \cdots + a_{m2}\mathbf{u}_m,$$
$$\vdots$$
$$\mathbf{v}_n = a_{1n}\mathbf{u}_1 + a_{2n}\mathbf{u}_2 + \cdots + a_{mn}\mathbf{u}_m.$$

Then, we let

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

Now we have

$$UA = V.$$

For the linear equations $V\mathbf{x} = \mathbf{0}$, we have

$$V\mathbf{x} = \mathbf{0} \qquad \Rightarrow \qquad UA\mathbf{x} = \mathbf{0}.$$

Because $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$ are independent, there is

$$V\mathbf{x} = \mathbf{0} \qquad \Rightarrow \qquad UA\mathbf{x} = \mathbf{0} \qquad \Rightarrow \qquad A\mathbf{x} = \mathbf{0}.$$

In addition, for the linear equations $A\mathbf{x} = \mathbf{0}$, we have

$$A\mathbf{x} = \mathbf{0} \qquad \Rightarrow \qquad UA\mathbf{x} = \mathbf{0} \qquad \Rightarrow \qquad V\mathbf{x} = \mathbf{0}.$$

Therefore, we have proved that the linear equations $V\mathbf{x} = \mathbf{0}$ and the linear equations $A\mathbf{x} = \mathbf{0}$ have the same solutions.

$$V\mathbf{x} = \mathbf{0} \qquad \Leftrightarrow \qquad A\mathbf{x} = \mathbf{0}.$$

Because $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are independent, the linear equations $V\mathbf{x} = \mathbf{0}$ can only have the solution $\mathbf{x} = \mathbf{0}$, there is

$$V\mathbf{x} = \mathbf{0} \quad \Leftrightarrow \quad \begin{cases} UA\mathbf{x} = \mathbf{0} & \Leftrightarrow \quad A\mathbf{x} = \mathbf{0} \\ \mathbf{x} = \mathbf{0} \end{cases}$$

Therefore, the linear equations $A\mathbf{x} = \mathbf{0}$ can only have the solution $\mathbf{x} = \mathbf{0}$.

$$A\mathbf{x} = \mathbf{0} \qquad \Leftrightarrow \qquad \mathbf{x} = \mathbf{0}.$$

However, because A is the $m \times n$ matrix and $m < n$, according to Gauss-Jordan Elimination, there must exist a free variable in RREF(A), and there must exist a solution for the linear equations $A\mathbf{x} = \mathbf{0}$ such that $\mathbf{x} \neq \mathbf{0}$, which makes a contradiction. Therefore, $m \geq n$.

■

(2) Because every vector in S_2 can be represented as a linear combination of the vectors in S_1, we have $m \geq n$. Because every vector in S_1 can be represented as a linear combination of the vectors in S_2, we have $m \leq n$. Therefore, $m = n$.

■

4. **Maximal Linearly Independent Subset**

(1)

- **Basis**

  $\forall$ a set of vectors with only one element $S = \{\mathbf{v}\}$. It is clear that S is its own maximal linearly independent subset.

- **Inductive Assumption**

  $\forall$ a set of vectors with i elements $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_i\}$, where $i \in \mathbb{N}^+$. We assume that it has a maximal linearly independent subset S'.

- **Inductive Step**

  $\forall$ a set of vectors with $i+1$ elements $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_i, \mathbf{v}_{i+1}\}$, where $i \in \mathbb{N}^+$. It is clear that the set $S_i = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_i\}$ is a subset of S. According to the inductive assumption, S_i has a maximal linearly independent subset S'. In addition, $S' \subset S$.

  - If $\mathbf{v}_{i+1}$ can be expressed as a linear combination of vectors in S', then S' is a maximal linearly independent subset of S.
  - If $\mathbf{v}_{i+1}$ can not be expressed as a linear combination of vectors in S', then $S' \cup \{\mathbf{v}_{i+1}\}$ is a maximal linearly independent subset of S.

■

(2) According to the definition of maximal linearly independent subset, $\mathbf{v}$ can be expressed as a linear combination of $\mathbf{v}_{k_1}, \mathbf{v}_{k_2}, \cdots, \mathbf{v}_{k_m}$. Assume that there exists

$$\mathbf{v} = \alpha_1 \mathbf{v}_{k_1} + \alpha_2 \mathbf{v}_{k_2} + \cdots + \alpha_n \mathbf{v}_{k_m}.$$

Then the equation

$$x_0 \mathbf{v} + x_1 \mathbf{v}_{k_1} + x_2 \mathbf{v}_{k_2} + \cdots + x_m \mathbf{v}_{k_m} = \mathbf{0}$$

has an nonzero solution

$$\mathbf{x} = \begin{bmatrix} 1 \\ -\alpha_1 \\ -\alpha_2 \\ \vdots \\ -\alpha_m \end{bmatrix}.$$

Therefore, $\mathbf{v}, \mathbf{v}_{k_1}, \mathbf{v}_{k_2}, \cdots, \mathbf{v}_{k_m}$ are linearly dependent.

■

(3) $\forall$ set of vectors S, assume that S_1 and S_2 are two maximal linearly independent subsets of S. According to **Core Property of Linear Independence**, because both S_1 and S_2 are linearly independent, and every vector in S_2 can be represented as a linear combination of the vectors in S_1; every vector in S_1 can be represented as a linear combination of the vectors in S_2. Therefore, S_1 and S_2 have the same number of vectors.

■

5. **Number Restriction and Rank**

Assume that a subset $S_u = \{\mathbf{u}_{u_1}, \mathbf{u}_{u_2}, \cdots, \mathbf{u}_{u_p}\}$ of S_1 is a maximal linearly independent subset of S_1; a subset $S_v = \{\mathbf{v}_{v_1}, \mathbf{v}_{v_2}, \cdots, \mathbf{v}_{v_q}\}$ of S_2 is a maximal linearly independent subset of S_2, where $1 \leq p \leq m, 1 \leq u_1 \leq u_2 \leq \cdots \leq u_p \leq m$, $1 \leq q \leq n, 1 \leq v_1 \leq v_2 \leq \cdots \leq v_q \leq n$ and $p, q, u_1, u_2, \cdots, u_p, v_1, v_2, \cdots, v_q \in \mathbb{N}^+$. Because every vector in S_2 can be represented as a linear combination of the vectors in S_1, every vector in S_v can be represented as a linear combination of the vectors in S_u. Then we have $q \leq p \leq m$.

- Assume that the vectors in S_2 is linearly independent, then we have $q = n \leq p \leq m$, which contradicts the condition that $m < n$. Therefore, the vectors in S_2 is linearly dependent.

■

- If the vectors in S_2 are linearly independent, then we have $q = n \leq p \leq m$.

■

- If the rank of S_1 is r_1 and the rank of S_2 is r_2, then we have $r_2 = q \leq p = r_1$.

■

6. **Basis and Dimension**

(1)

- According to the definition of basis, $\mathbf{u}$ can be represented as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$. Assume that there exists

$$\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n,$$

and

$$\mathbf{u} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_n\mathbf{v}_n,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_n \in \mathbb{R}$. Then there is

$$(\alpha_1 - \beta_1)\mathbf{v}_1 + (\alpha_2 - \beta_2)\mathbf{v}_2 + \cdots + (\alpha_n - \beta_n)\mathbf{v}_n = \mathbf{0}.$$

Because the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent, the only solution of this equation is

$$\alpha_1 - \beta_1 = 0, \quad \alpha_2 - \beta_2 = 0, \quad \cdots, \quad \alpha_n - \beta_n = 0.$$

Hence

$$\alpha_1 = \beta_1, \quad \alpha_2 = \beta_2, \quad \cdots, \quad \alpha_n = \beta_n.$$

■

- According to the definition of basis, $\mathbf{u}$ can be represented as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$. Assume that there exists

$$\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n.$$

Then the equation

$$x_0\mathbf{u} + x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \mathbf{0}$$

has an nonzero solution

$$\mathbf{x} = \begin{bmatrix} 1 \\ -\alpha_1 \\ -\alpha_2 \\ \vdots \\ -\alpha_n \end{bmatrix}.$$

Therefore, $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

■

(2) According to **Core Property of Linear Independence**, because both S_1 and S_2 are linearly independent, and every vector in S_2 can be represented as a linear combination of the vectors in S_1; every vector in S_1 can be represented as a linear combination of the vectors in S_2. Therefore, S_1 and S_2 have the same number of vectors, $m = n$.

(3)
- $\forall\ n \in \mathbb{N}^+$, assume that the identity matrix $I_{n\times n}$ is

$$I = I_{n\times n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \\ | & | & \cdots & | \end{bmatrix}.$$

  Then the vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ form a basis of $\mathbf{R}^n$, because $\forall$ vector $\mathbf{v} \in \mathbf{R}^n$ such that

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

  can be expressed as a linear comination of $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$:

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n,$$

  where $v_1, v_2, \cdots, v_n \in \mathbb{R}$. Therefore, $\forall\ n \in \mathbb{N}^+$, the dimension of $\mathbf{R}^n$ is n.

■

- Assume that there is a set of vectors S is a basis of $\mathbf{R}^m$, so $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ can be expressed as a linear combination of the vectors in S. Because $m < n$, according to the first item of **Number Restriction and Rank**, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

■

- Assume that a set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is a set of n independent vectors in $\mathbf{R}^n$. Assume that there exists a vector $\mathbf{v}$ which can not be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$. Therefore, the set of vectors $S' = S \cup \{\mathbf{v}\} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n, \mathbf{v}\}$ is linearly independent. However, the dimension of $\mathbf{R}^n$ is n, so we can find a basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$

of $\mathbf{R}^n$. Thus, the vectors in S' can be expressed as linear combinations of vectors in E. In addition, both vectors in S' and vectors in E are linearly independent. Therefore, the rank of S' is no more than the rank of E, which means that $n+1 \leq n$, which is a contradiction. Therefore, $\forall\ n$ independent vectors in $\mathbf{R}^n$ must span $\mathbf{R}^n$, which means that they are a basis.

■

(4) $\forall\ m \times n$ matrix A where $m, n \in \mathbb{N}^+$ such that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix}.$$

Assume that the dimension of the vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_m$ is r, where $1 \leq r \leq m$ and $r \in \mathbb{N}^+$.

i. The proof that switching rows does not change the dimension of the vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m$.

$\forall\ 1 \leq i < j \leq m$ and $i, j \in \mathbb{N}^+$, the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}, \cdots, \mathbf{r}_{j-1}, \mathbf{r}_j, \mathbf{r}_{j+1}, \cdots \mathbf{r}_m$$

are the same as the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, \mathbf{r}_j, \mathbf{r}_{i+1}, \cdots, \mathbf{r}_{j-1}, \mathbf{r}_i, \mathbf{r}_{j+1}, \cdots \mathbf{r}_m.$$

Therefore, these two sets of vectors have the same dimension.

ii. The proof that multiplying a row by a number does not change the dimension of the vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m$.

$\forall\ 1 \leq i \leq m$ and $i \in \mathbb{N}^+$, assume that the row vector $\mathbf{r}_i$ is multiplied by a number k, where $k \in \mathbb{R}$ and $k \neq 0$. Because

$$k\mathbf{r}_i = k \cdot \mathbf{r}_i, \qquad \mathbf{r}_i = \frac{1}{k} \cdot (k\mathbf{r}_i),$$

then the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}, \cdots \mathbf{r}_m$$

and the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, k\mathbf{r}_i, \mathbf{r}_{i+1}, \cdots \mathbf{r}_m.$$

can be represented as a linear combination with each other. Assume that the dimension of the second set vectors is r'. According to the proof of the part 1, we can get

$$r \leq r', \qquad r \geq r' \qquad \Rightarrow \qquad r = r'.$$

iii. The proof that adding multiples of rows does not change the dimension of the vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m$.
$\forall\ 1 \leq i < j \leq m$ and $i, j \in \mathbb{N}^+$,

- Assume that the row vector $\mathbf{r}_i$ is added by the vector $\mathbf{r}_j$ with a multiple of k, where $k \in \mathbb{R}$ and $k \neq 0$. Because

$$\mathbf{r}_i + k\mathbf{r}_j = \mathbf{r}_i + k \cdot \mathbf{r}_j, \qquad \mathbf{r}_i = (\mathbf{r}_i + k\mathbf{r}_j) - k \cdot \mathbf{r}_j,$$

then the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}, \cdots, \mathbf{r}_{j-1}, \mathbf{r}_j, \mathbf{r}_{j+1}, \cdots \mathbf{r}_m$$

and the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, \mathbf{r}_i + k\mathbf{r}_j, \mathbf{r}_{i+1}, \cdots, \mathbf{r}_{j-1}, \mathbf{r}_j, \mathbf{r}_{j+1}, \cdots \mathbf{r}_m$$

can be represented as a linear combination with each other. Assume that the dimension of the second set vectors is r'. According to the proof of the part 1, we can get

$$r \leq r', \qquad r \geq r' \qquad \Rightarrow \qquad r = r'.$$

- Assume that the row vector $\mathbf{r}_j$ is added by the vector $\mathbf{r}_i$ with a multiple of k, where $k \in \mathbb{R}$ and $k \neq 0$. Because

$$\mathbf{r}_j + k\mathbf{r}_i = \mathbf{r}_j + k \cdot \mathbf{r}_i, \qquad \mathbf{r}_j = (\mathbf{r}_j + k\mathbf{r}_i) - k \cdot \mathbf{r}_i,$$

then the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}, \cdots, \mathbf{r}_{j-1}, \mathbf{r}_j, \mathbf{r}_{j+1}, \cdots \mathbf{r}_m$$

and the vectors

$$\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}, \cdots, \mathbf{r}_{j-1}, \mathbf{r}_j + k\mathbf{r}_i, \mathbf{r}_{j+1}, \cdots \mathbf{r}_m$$

can be represented as a linear combination with each other. Assume that the dimension of the second set vectors is r'. According to the proof of the part 1, we can get

$$r \leq r', \qquad r \geq r' \qquad \Rightarrow \qquad r = r'.$$

To sum up, the row operations do not change the dimension of the row space of matrices.

■

Theorem A.11 (Fundamental Theorem of Linear Algebra, Part 1) $\forall\, m \times n$ matrix A where $m, n \in \mathbb{N}^+$, suppose that the rank of A is r $(1 \leq r \leq \min\{m, n\})$. The column space$\boldsymbol{C}(A)$ and row space $\boldsymbol{C}(A^{\mathrm{T}})$ both have dimension r. The nullspace $\boldsymbol{N}(A)$ has dimension $n - r$. The left nullspace $\boldsymbol{N}(A^{\mathrm{T}})$ has dimension $m - r$.

Proof.

1. **The proof that the rank of A is the same as the dimension of the row space of A.**

   Because the row operations of matrices do not change the dimension of row space of matrices, the dimension of the row space of A is the same as the dimension of the row space of the RREF of A. It is clear that the dimension of the row space of RREF(A) is the number of pivots rows of RREF(A). Therefore, the dimension of the row space of RREF(A) is the same as the rank of RREF(A), which is also the rank of A.

2. **The proof that the dimension of the column space of A is the same as the dimension of the row space of A.**

   Assume that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix}.$$

Assume that the dimension of row space of A is r, the dimension of column space of A is c, where $1 \leq r \leq m, 1 \leq c \leq n$ and $r, c \in \mathbb{N}^+$.

(1) The proof of $r \leq c$.

Suppose that the vectors $\mathbf{c}_{i_1}, \mathbf{c}_{i_2}, \cdots, \mathbf{c}_{i_c}$ form a basis of the column space of A, where $1 \leq i_1 \leq i_2 \leq \cdots \leq i_c \leq n$. Then, suppose that

$$\begin{aligned}
\mathbf{c}_1 &= c_{11}\mathbf{c}_{i_1} + c_{21}\mathbf{c}_{i_2} + \cdots + c_{c1}\mathbf{c}_{i_c}, \\
\mathbf{c}_2 &= c_{12}\mathbf{c}_{i_1} + c_{22}\mathbf{c}_{i_2} + \cdots + c_{c2}\mathbf{c}_{i_c}, \\
&\vdots \\
\mathbf{c}_n &= c_{1n}\mathbf{c}_{i_1} + c_{2n}\mathbf{c}_{i_2} + \cdots + c_{cn}\mathbf{c}_{i_c}.
\end{aligned}$$

Let

$$C = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_{i_1} & \mathbf{c}_{i_2} & \cdots & \mathbf{c}_{i_c} \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} a_{1i_1} & a_{1i_2} & \cdots & a_{1i_c} \\ a_{2i_1} & a_{2i_2} & \cdots & a_{2i_c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mi_1} & a_{mi_2} & \cdots & a_{mi_c} \end{bmatrix}$$

and

$$C' = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{c1} & c_{c2} & \cdots & c_{cn} \end{bmatrix} = \begin{bmatrix} — & \mathbf{r}_{C'_1}^{\mathrm{T}} & — \\ — & \mathbf{r}_{C'_2}^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_{C'_c}^{\mathrm{T}} & — \end{bmatrix}.$$

It is clear that C is a $m \times c$ matrix and C' is a $c \times n$ matrix. Because

$$\begin{aligned}
A &= \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix} \\
&= \begin{bmatrix} c_{11}\mathbf{c}_{i_1} & c_{12}\mathbf{c}_{i_1} & & c_{1n}\mathbf{c}_{i_1} \\ + & + & & + \\ c_{21}\mathbf{c}_{i_2} & c_{22}\mathbf{c}_{i_2} & & c_{2n}\mathbf{c}_{i_2} \\ + & + & & + \\ \vdots & \vdots & \cdots & \vdots \\ + & + & & + \\ c_{c1}\mathbf{c}_{i_c} & c_{c2}\mathbf{c}_{i_c} & & c_{cn}\mathbf{c}_{i_c} \end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_{i_1} & \mathbf{c}_{i_2} & \cdots & \mathbf{c}_{i_c} \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{c1} & c_{c2} & \cdots & c_{cn} \end{bmatrix}$$

$$= CC'.$$

We have

$$A = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix}$$

$$= CC'$$

$$= \begin{bmatrix} a_{1i_1} & a_{1i_2} & \cdots & a_{1i_c} \\ a_{2i_1} & a_{2i_2} & \cdots & a_{2i_c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mi_1} & a_{mi_2} & \cdots & a_{mi_c} \end{bmatrix} \begin{bmatrix} — & \mathbf{r}_{C_1'}^{\mathrm{T}} & — \\ — & \mathbf{r}_{C_2'}^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_{C_c'}^{\mathrm{T}} & — \end{bmatrix}$$

$$= \begin{bmatrix} a_{1i_1}\mathbf{r}_{C_1'}^{\mathrm{T}} + a_{1i_2}\mathbf{r}_{C_2'}^{\mathrm{T}} + \cdots + a_{1i_c}\mathbf{r}_{C_c'}^{\mathrm{T}} \\ a_{2i_1}\mathbf{r}_{C_1'}^{\mathrm{T}} + a_{2i_2}\mathbf{r}_{C_2'}^{\mathrm{T}} + \cdots + a_{2i_c}\mathbf{r}_{C_c'}^{\mathrm{T}} \\ \vdots \\ a_{mi_1}\mathbf{r}_{C_1'}^{\mathrm{T}} + a_{mi_2}\mathbf{r}_{C_2'}^{\mathrm{T}} + \cdots + a_{mi_c}\mathbf{r}_{C_c'}^{\mathrm{T}}. \end{bmatrix}.$$

Therefore, the vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m$ can be represented as a linear combination of the vectors $\mathbf{r}_{C_1'}, \mathbf{r}_{C_2'}, \cdots, \mathbf{r}_{C_c'}$, so the dimension of the vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m$ is not more than the dimension of the vectors $\mathbf{r}_{C_1'}, \mathbf{r}_{C_2'}, \cdots, \mathbf{r}_{C_c'}$. Assume that the dimension of the vectors $\mathbf{r}_{C_1'}, \mathbf{r}_{C_2'}, \cdots, \mathbf{r}_{C_c'}$ is $r_{C'}$. Then we have

$$r \leq r_{C'} \leq c.$$

(2) The proof of $c \leq r$.

Suppose that the vectors $\mathbf{r}_{j_1}, \mathbf{r}_{j_2}, \cdots, \mathbf{r}_{j_r}$ form a basis of the row space of A, where $1 \leq j_1 \leq j_2 \leq \cdots \leq j_r \leq m$. Then, suppose that

$$\mathbf{r}_1 = r_{11}\mathbf{r}_{j_1} + r_{12}\mathbf{r}_{j_2} + \cdots + r_{1r}\mathbf{r}_{j_r},$$

$$\mathbf{r}_2 = r_{21}\mathbf{r}_{j_1} + r_{22}\mathbf{r}_{j_2} + \cdots + r_{2r}\mathbf{r}_{j_r},$$

$$\vdots$$

$$\mathbf{r}_m = r_{m1}\mathbf{r}_{j_1} + r_{m2}\mathbf{r}_{j_2} + \cdots + r_{mr}\mathbf{r}_{j_r}.$$

Let

$$R = \begin{bmatrix} — & \mathbf{r}_{j_1}^{\mathrm{T}} & — \\ — & \mathbf{r}_{j_2}^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_{j_r}^{\mathrm{T}} & — \end{bmatrix} = \begin{bmatrix} a_{j_1 1} & a_{j_1 2} & \cdots & a_{j_1 n} \\ a_{j_2 1} & a_{j_2 2} & \cdots & a_{j_2 n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j_r 1} & a_{j_r 2} & \cdots & a_{j_r n} \end{bmatrix}$$

and

$$R' = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1r} \\ r_{21} & r_{22} & \cdots & r_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mr} \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_{R'_1} & \mathbf{c}_{R'_2} & \cdots & \mathbf{c}_{R'_r} \\ | & | & \cdots & | \end{bmatrix}.$$

It is clear that R' is a $m \times r$ matrix and R is a $r \times n$ matrix. Because

$$\begin{aligned} A &= \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix} \\ &= \begin{bmatrix} r_{11}\mathbf{r}_{j_1} + r_{12}\mathbf{r}_{j_2} + \cdots + r_{1r}\mathbf{r}_{j_r} \\ r_{21}\mathbf{r}_{j_1} + r_{22}\mathbf{r}_{j_2} + \cdots + r_{2r}\mathbf{r}_{j_r} \\ \vdots \\ r_{m1}\mathbf{r}_{j_1} + r_{m2}\mathbf{r}_{j_2} + \cdots + r_{mr}\mathbf{r}_{j_r} \end{bmatrix} \\ &= \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1r} \\ r_{21} & r_{22} & \cdots & r_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mr} \end{bmatrix} \begin{bmatrix} — & \mathbf{r}_{j_1}^{\mathrm{T}} & — \\ — & \mathbf{r}_{j_2}^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_{j_r}^{\mathrm{T}} & — \end{bmatrix} \\ &= R'R. \end{aligned}$$

We have

$$A = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix}$$

$$
\begin{aligned}
&= R'R \\
&= \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_{R'_1} & \mathbf{c}_{R'_2} & \cdots & \mathbf{c}_{R'_r} \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} a_{j_1 1} & a_{j_1 2} & \cdots & a_{j_1 n} \\ a_{j_2 1} & a_{j_2 2} & \cdots & a_{j_2 n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j_r 1} & a_{j_r 2} & \cdots & a_{j_r n} \end{bmatrix} \\
&= \begin{bmatrix} a_{j_1 1}\mathbf{c}_{R'_1} & a_{j_1 2}\mathbf{c}_{R'_1} & & a_{j_1 n}\mathbf{c}_{R'_1} \\ + & + & & + \\ a_{j_2 1}\mathbf{c}_{R'_2} & a_{j_2 2}\mathbf{c}_{R'_2} & & a_{j_2 n}\mathbf{c}_{R'_2} \\ + & + & & + \\ \vdots & \vdots & \cdots & \vdots \\ + & + & & + \\ a_{j_r 1}\mathbf{c}_{R'_r} & a_{j_r 2}\mathbf{c}_{R'_r} & & a_{j_r n}\mathbf{c}_{R'_r} \end{bmatrix}.
\end{aligned}
$$

Therefore, the vectors $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n$ can be represented as a linear combination of the vectors $\mathbf{c}_{R'_1}, \mathbf{c}_{R'_2}, \cdots, \mathbf{c}_{R'_r}$, so the dimension of the vectors $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n$ is not more than the dimension of the vectors $\mathbf{c}_{R'_1}, \mathbf{c}_{R'_2}, \cdots, \mathbf{c}_{R'_r}$. Assume that the dimension of the vectors $\mathbf{c}_{R'_1}, \mathbf{c}_{R'_2}, \cdots, \mathbf{c}_{R'_r}$ is $c_{R'}$. Then we have

$$c \leq c_{R'} \leq r.$$

To sum up, $r = c$.

3. **The proof that the nullspace $\boldsymbol{N}(A)$ has dimension $n - r$ and the left nullspace $\boldsymbol{N}(A^{\mathrm{T}})$ has dimension $m - r$.**

   According to the special solution to $A\mathbf{x} = 0$, it can be expressed as a linear combination of a series of $n-$dimensional vectors, and these vectors are also independent. Therefore, they form a basis of $\boldsymbol{N}(A)$. Because the number of these vectors is $n - r$, the nullspace $\boldsymbol{N}(A)$ has dimension $n - r$. With the same principle, the left nullspace $\boldsymbol{N}(A^{\mathrm{T}})$ has dimension $m - r$.

■

Theorem A.12 (Properties of Orthogonal Subspaces)

1. **Independence**

$\forall$ two orthogonal subspaces $\boldsymbol{U}$ and $\boldsymbol{V}$, assume that a set of vectors $S_u = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$ is a basis of $\boldsymbol{U}$ and a set of vectors $S_v = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is a basis of $\boldsymbol{V}$, where $m, n \in \mathbb{N}^+$. Then the vectors in the set $S = S_u \cup S_v = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ are linearly independent.

2. **Symmetry**

   $\forall$ subspaces $\boldsymbol{V}$, its orthogonal complement is $\boldsymbol{V}^\perp$. Then the orthogonal complement of $\boldsymbol{V}^\perp$ is $\boldsymbol{V}$, which means that $(\boldsymbol{V}^\perp)^\perp = \boldsymbol{V}$.

3. **Complementarity**

   $\forall$ subspaces $\boldsymbol{V}$, its orthogonal complement is $\boldsymbol{V}^\perp$. Assume that the vectors in $\boldsymbol{V}$ and $\boldsymbol{V}^\perp$ are $n-$dimensional, where $h \in \mathbb{N}^+$. Then if a set of vectors S_v is a basis of $\boldsymbol{V}$ and a set of vectors $S_{v^\perp}$ is a basis of $\boldsymbol{V}^\perp$, then the set of vectors $S = S_v \cup S_{v^\perp}$ is a basis of $\mathbf{R}^n$.

Proof.

1. Assume that the vectors in the set $S = S_u \cup S_v = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ are linearly dependent.

   - Assume that there is a vector $\mathbf{u}_i \in S_u$ can be expressed as a linear combination of the vectors in the set $S - \{\mathbf{u}_i\}$, where $1 \leq i \leq m$ and $i \in \mathbb{N}^+$. Let

     $$\mathbf{u}_i = a_1\mathbf{u}_1 + \cdots + a_{i-1}\mathbf{u}_{i-1} + a_{i+1}\mathbf{u}_{i+1} + \cdots + a_m\mathbf{u}_m + b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n,$$

     where $a_1, \cdots, a_m, b_1, \cdots, b_n \in \mathbb{R}$. Then we have

     $$-a_1\mathbf{u}_1 - \cdots - a_{i-1}\mathbf{u}_{i-1} + \mathbf{u}_i - a_{i+1}\mathbf{u}_{i+1} - \cdots - a_m\mathbf{u}_m = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n.$$

     Let

     $$\mathbf{u} = -a_1\mathbf{u}_1 - \cdots - a_{i-1}\mathbf{u}_{i-1} + \mathbf{u}_i - a_{i+1}\mathbf{u}_{i+1} - \cdots - a_m\mathbf{u}_m.$$

     Because $\boldsymbol{U}$ is orthogonal to $\boldsymbol{V}$, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$ are orthogonal to $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$. There is

     $$\mathbf{u}^{\mathrm{T}}\mathbf{u} = \mathbf{u}^{\mathrm{T}}(b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n) = b_1\mathbf{u}^{\mathrm{T}}\mathbf{v}_1 + \cdots + b_n\mathbf{u}^{\mathrm{T}}\mathbf{v}_n = 0.$$

     Thus,

     $$\mathbf{u} = -a_1\mathbf{u}_1 - \cdots - a_{i-1}\mathbf{u}_{i-1} + \mathbf{u}_i - a_{i+1}\mathbf{u}_{i+1} - \cdots - a_m\mathbf{u}_m = \mathbf{0},$$

$$\mathbf{u}_i = a_1\mathbf{u}_1 + \cdots + a_{i-1}\mathbf{u}_{i-1} + a_{i+1}\mathbf{u}_{i+1} + \cdots + a_m\mathbf{u}_m.$$

Because the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$ are linearly independent, there is

$$a_1 = \cdots = a_{i-1} = a_{i+1} = \cdots = a_m = 0.$$

Therefore,

$$\mathbf{u}_i = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n,$$
$$\mathbf{u}_i^{\mathrm{T}}\mathbf{u}_i = \mathbf{u}_i^{\mathrm{T}}(b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n) = b_1\mathbf{u}^{\mathrm{T}}\mathbf{v}_1 + \cdots + b_n\mathbf{u}^{\mathrm{T}}\mathbf{v}_n = 0.$$

Now we get $\mathbf{u}_i = \mathbf{0}$, which can not be a part of basis, so this is a cotradiction.

- Assume that there is a vector $\mathbf{v}_j \in S_v$ can be expressed as a linear combination of the vectors in the set $S - \{\mathbf{v}_j\}$, where $1 \leq j \leq n$ and $j \in \mathbb{N}^+$. Let

$$\mathbf{v}_j = a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m + b_1\mathbf{v}_1 + \cdots + b_{j-1}\mathbf{v}_{j-1} + b_{j+1}\mathbf{v}_{j+1} + \cdots + b_n\mathbf{v}_n,$$

where $a_1, \cdots, a_m, b_1, \cdots, b_n \in \mathbb{R}$. Then we have

$$-b_1\mathbf{v}_1 - \cdots - b_{j-1}\mathbf{v}_{j-1} + \mathbf{v}_j - b_{j+1}\mathbf{v}_{j+1} - \cdots - b_n\mathbf{v}_n = a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m.$$

Let

$$\mathbf{v} = -b_1\mathbf{v}_1 - \cdots - b_{j-1}\mathbf{v}_{j-1} + \mathbf{v}_j - b_{j+1}\mathbf{v}_{j+1} - \cdots - b_n\mathbf{v}_n.$$

Because $\boldsymbol{U}$ is orthogonal to $\boldsymbol{V}$, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$ are orthogonal to $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$. There is

$$\mathbf{v}^{\mathrm{T}}\mathbf{v} = \mathbf{v}^{\mathrm{T}}(a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m) = a_1\mathbf{v}^{\mathrm{T}}\mathbf{u}_1 + \cdots + a_m\mathbf{v}^{\mathrm{T}}\mathbf{u}_m = 0.$$

Thus,

$$\mathbf{v} = -b_1\mathbf{v}_1 - \cdots - b_{j-1}\mathbf{v}_{j-1} + \mathbf{v}_j - b_{j+1}\mathbf{v}_{j+1} - \cdots - b_n\mathbf{v}_n. = \mathbf{0},$$
$$\mathbf{v}_j = b_1\mathbf{v}_1 + \cdots + b_{j-1}\mathbf{v}_{j-1} + b_{j+1}\mathbf{v}_{j+1} + \cdots + b_n\mathbf{v}_n.$$

Because the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent, there is

$$b_1 = \cdots = b_{j-1} = b_{j+1} = \cdots = b_n = 0.$$

Therefore,

$$\mathbf{v}_j = a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m,$$
$$\mathbf{v}_j^{\mathrm{T}}\mathbf{v}_j = \mathbf{v}_j^{\mathrm{T}}(a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m) = a_1\mathbf{v}^{\mathrm{T}}\mathbf{u}_1 + \cdots + a_m\mathbf{v}^{\mathrm{T}}\mathbf{u}_m = 0.$$

Now we get $\mathbf{v}_j = \mathbf{0}$, which can not be a part of basis, so this is a cotradiction.

To sum up, every vector in S can not be expressed as a linear combination of other vectors in S, so the vectors in the set $S = S_u \cup S_v = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ are linearly dependent.

■

2. Assume that the orthogonal complement of $\boldsymbol{V}^\perp$ is not $\boldsymbol{V}$, which means that there exists a vector $\mathbf{r}$ such that $\mathbf{r}$ is orthogonal to $\boldsymbol{V}$ but $\mathbf{r} \notin \boldsymbol{V}^\perp$. Assume that a set of vectors $S = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m\}$ is a basis of $\boldsymbol{V}^\perp$; a set of vectors $S' = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$ is a basis of $\boldsymbol{V}$, where $m, n \in \mathbb{N}^+$. Assume that the vectors in $\boldsymbol{V}$ and $\boldsymbol{V}^\perp$ are $k-$dimensional, where $k \geq \max\{m, n\}$ and $k \in \mathbb{N}^+$. Let

$$A = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix}, \qquad A' = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \\ — & \mathbf{r}^{\mathrm{T}} & — \end{bmatrix}.$$

(1) **The proof that $\boldsymbol{V} = \boldsymbol{N}(\boldsymbol{A})$**

Because $\boldsymbol{V}$ is orthogonal to $\boldsymbol{V}^\perp$, the vectors in $S = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m\}$ are orthogonal to the vectors in $S' = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$. There is

$$A\mathbf{x}_1 = A\mathbf{x}_2 = \cdots = A\mathbf{x}_n = \mathbf{0}.$$

$\forall$ $k-$dimensional vector $\mathbf{x}$,

- If $\mathbf{x} \in \boldsymbol{V}$, then it can be expressed as a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$. Let

  $$\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n,$$

  where $a_1, a_2, \cdots, a_n \in \mathbb{R}$. Thus,

  $$A\mathbf{x} = A(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n) = a_1A\mathbf{x}_1 + a_2A\mathbf{x}_2 + \cdots + a_nA\mathbf{x}_n = \mathbf{0}.$$

  Therefore, $\mathbf{x} \in \boldsymbol{N}(A)$.

- If $\mathbf{x} \in \boldsymbol{N}(A)$, then there is

  $$A\mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix} \mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ — & \mathbf{r}_2^{\mathrm{T}}\mathbf{x} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}}\mathbf{x} & — \end{bmatrix} = \mathbf{0}.$$

Thus,

$$\mathbf{r}_1^{\mathrm{T}}\mathbf{x} = \mathbf{r}_2^{\mathrm{T}}\mathbf{x} = \cdots = \mathbf{r}_m^{\mathrm{T}}\mathbf{x} = 0.$$

$\forall$ vector $\mathbf{v} \in \boldsymbol{V}^{\perp}$, it can be expressed as a linear combination of $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m$. Let

$$\mathbf{v} = b_1\mathbf{r}_1 + b_2\mathbf{r}_2 + \cdots + b_m\mathbf{r}_m,$$

where $b_1, b_2, \cdots, b_m \in \mathbb{R}$. Thus,

$$\mathbf{v}^{\mathrm{T}}\mathbf{x} = (b_1\mathbf{r}_1 + b_2\mathbf{r}_2 + \cdots + b_m\mathbf{r}_n)^{\mathrm{T}}\mathbf{x} = b_1\mathbf{r}_1^{\mathrm{T}}\mathbf{x} + b_2\mathbf{r}_2^{\mathrm{T}}\mathbf{x} + \cdots + b_m\mathbf{r}_m^{\mathrm{T}}\mathbf{x} = 0.$$

Therefore, $\mathbf{x}$ is orthogonal to $\boldsymbol{V}^{\perp}$, which means that $\mathbf{x} \in \boldsymbol{V}$.

To sum up, $\boldsymbol{V} = \boldsymbol{N}(\boldsymbol{A})$. In addition, $k = m + n$.

(2) **The proof that $\boldsymbol{N}(\boldsymbol{A}) = \boldsymbol{N}(\boldsymbol{A'})$**

$\forall$ $k-$dimensional vector $\mathbf{x}$

- If $\mathbf{x} \in \boldsymbol{N}(A)$, then

$$A\mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix} \mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ — & \mathbf{r}_2^{\mathrm{T}}\mathbf{x} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}}\mathbf{x} & — \end{bmatrix} = \mathbf{0}.$$

Because $\mathbf{r}$ is orthogonal to $\boldsymbol{V}$, which means that $\mathbf{r}$ is orthogonal to $\boldsymbol{N}(A)$, $\mathbf{r}^{\mathrm{T}}\mathbf{x} = 0$. Now

$$A'\mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \\ — & \mathbf{r}^{\mathrm{T}} & — \end{bmatrix} \mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}}\mathbf{x} & — \\ — & \mathbf{r}^{\mathrm{T}}\mathbf{x} & — \end{bmatrix} = \mathbf{0}.$$

Therefore, $\mathbf{x} \in \boldsymbol{N}(A')$.

- If $\mathbf{x} \in \boldsymbol{N}(A')$, then

$$A'\mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \\ — & \mathbf{r}^{\mathrm{T}} & — \end{bmatrix} \mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}}\mathbf{x} & — \\ — & \mathbf{r}^{\mathrm{T}}\mathbf{x} & — \end{bmatrix} = \mathbf{0}.$$

Thus,

$$\mathbf{r}_1^{\mathrm{T}}\mathbf{x} = \mathbf{r}_2^{\mathrm{T}}\mathbf{x} = \cdots = \mathbf{r}_m^{\mathrm{T}}\mathbf{x} = 0.$$

Then

$$A\mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix} \mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ — & \mathbf{r}_2^{\mathrm{T}}\mathbf{x} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}}\mathbf{x} & — \end{bmatrix} = \mathbf{0}.$$

Therefore, $\mathbf{x} \in \boldsymbol{N}(A)$.

To sum up, $\boldsymbol{N(A)} = \boldsymbol{N(A')}$.

(3) **The contradiction**

Because $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_m$ are linearly independent, the dimension of $\boldsymbol{C}(A^{\mathrm{T}})$ is m, and the dimension of $\boldsymbol{N}(A)$ is $k - m = n$. However, the dimension of $\boldsymbol{C}(A'^{\mathrm{T}})$ is $m + 1$, because $\mathbf{r}$ is a linear independent vector compared with vectors $\mathbf{r}_1, \cdots, \mathbf{r}_m$. Thus the dimension of $\boldsymbol{N}(A_1)$ is $k - (m + 1) = n - 1$, which cotradicts the conclusion that $\boldsymbol{N}(A) = \boldsymbol{N}(A')$.

To sum up, there does not exist a vector $\mathbf{r}$ such that $\mathbf{r}$ is orthogonal to $\boldsymbol{V}$ but $\mathbf{r} \notin \boldsymbol{V}^{\perp}$. Therefore, $\forall$ subspaces $\boldsymbol{V}$, its orthogonal complement is $\boldsymbol{V}^{\perp}$. Then the orthogonal complement of $\boldsymbol{V}^{\perp}$ is $\boldsymbol{V}$, which means that $(\boldsymbol{V}^{\perp})^{\perp} = \boldsymbol{V}$.

■

3. Assume that $S_v = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_r\}$ is a basis of $\boldsymbol{V}$; a set of vectors $S_{v^{\perp}} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p\}$ is a basis of $\boldsymbol{V}^{\perp}$, where $1 \leq p, r \leq n$ and $p, r \in \mathbb{N}^+$. Let

$$A = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_r^{\mathrm{T}} & — \end{bmatrix}.$$

According to the proof of Symmetry, there is $\boldsymbol{V}^{\perp} = \boldsymbol{N}(A)$. The dimension of $\boldsymbol{V}$ is r, and the dimension of $\boldsymbol{V}^{\perp}$ is p. Thus, we have

$$r + p = n.$$

According to the proof of Independence, the vectors in the set $S = S_v \cup S_{v^\perp}$ are linearly independent. Therefore, these n linearly independent vectors can form a basis of $\mathbf{R}^n$.

■

Theorem A.13 (Fundamental Theorem of Linear Algebra, Part 2) $\forall\, m \times n$ matrix A where $m, n \in \mathbb{N}^+$,

1. $\boldsymbol{N}(A)$ is the orthogonal complement of $\boldsymbol{C}(A^{\mathrm{T}})$, which is in $\mathbf{R}^n$.
2. $\boldsymbol{C}(A^{\mathrm{T}})$ is the orthogonal complement of $\boldsymbol{N}(A)$, which is in $\mathbf{R}^n$.
3. $\boldsymbol{N}(A^{\mathrm{T}})$ is the orthogonal complement of $\boldsymbol{C}(A)$, which is in $\mathbf{R}^m$.
4. $\boldsymbol{C}(A)$ is the orthogonal complement of $\boldsymbol{N}(A^{\mathrm{T}})$, which is in $\mathbf{R}^m$.

Proof. Assume that

$$A = \begin{bmatrix} | & \cdots & | \\ \mathbf{c}_1 & \cdots & \mathbf{c}_n \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix}.$$

1. $\forall$ n-dimensional vector $\mathbf{x}$ which is orthogonal to $\boldsymbol{C}(A^{\mathrm{T}})$, there is

   $$\mathbf{r}_1^{\mathrm{T}}\mathbf{x} = \cdots = \mathbf{r}_m^{\mathrm{T}}\mathbf{x} = 0.$$

   Thus,

   $$A\mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix} \mathbf{x} = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}}\mathbf{x} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}}\mathbf{x} & — \end{bmatrix} = \mathbf{0}.$$

   Therefore, $\mathbf{x} \in \boldsymbol{N}(A)$, $\boldsymbol{N}(A)$ is the orthogonal complement of $\boldsymbol{C}(A^{\mathrm{T}})$ (in $\mathbf{R}^n$).

2. Because $\boldsymbol{N}(A)$ is the orthogonal complement of $\boldsymbol{C}(A^{\mathrm{T}})$, which is in $\mathbf{R}^n$, $\boldsymbol{C}(A^{\mathrm{T}})$ is also the orthogonal complement of $\boldsymbol{N}(A)$, which is in $\mathbf{R}^n$.

3. $\forall$ m-dimensional vector $\mathbf{x}$ which is orthogonal to $\boldsymbol{C}(A)$, there is

   $$\mathbf{c}_1^{\mathrm{T}}\mathbf{x} = \cdots = \mathbf{c}_n^{\mathrm{T}}\mathbf{x} = 0.$$

Thus,

$$A^{\mathrm{T}}\mathbf{x} = \begin{bmatrix} — & \mathbf{c}_1^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{c}_n^{\mathrm{T}} & — \end{bmatrix} \mathbf{v} = \begin{bmatrix} — & \mathbf{c}_1^{\mathrm{T}}\mathbf{x} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{c}_n^{\mathrm{T}}\mathbf{x} & — \end{bmatrix} = \mathbf{0}.$$

Therefore, $\mathbf{x} \in \boldsymbol{N}(A^{\mathrm{T}})$, $\boldsymbol{N}(A^{\mathrm{T}})$ is the orthogonal complement of $\boldsymbol{C}(A)$ (in $\mathbf{R}^m$).

4. Because $\boldsymbol{N}(A^{\mathrm{T}})$ is the orthogonal complement of $\boldsymbol{C}(A)$, which is in $\mathbf{R}^m$, $\boldsymbol{C}(A)$ is also the orthogonal complement of $\boldsymbol{N}(A^{\mathrm{T}})$, which is in $\mathbf{R}^m$.

■

Theorem A.14 (The Decomposition of The Solution to $A\mathbf{x} = \mathbf{b}$) $\forall$ $m \times n$ matrix A where $m, n \in \mathbb{N}^+$, $\forall$ vector $\mathbf{b} \in \boldsymbol{C}(A)$ and $\mathbf{b} \neq \mathbf{0}$. For the linear equations $A\mathbf{x} = \mathbf{b}$:

1. $\exists$ a unique $n-$dimensional vector $\mathbf{x}_r \in \boldsymbol{C}(A^{\mathrm{T}})$ such that $A\mathbf{x}_r = \mathbf{b}$.

2. $\forall$ solution $\mathbf{x}_c$ to $A\mathbf{x} = \mathbf{0}$ can be decomposed as $\mathbf{x}_c = \mathbf{x}_r + \mathbf{x}_n$, where $\mathbf{x}_n \in \boldsymbol{N}(A)$.

Proof. Assume that

$$A = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix},$$

the rank of A is r, where $1 \leq r \leq \min\{m, n\}$; a set of vectors $S_r = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_r\}$ is a basis of $\boldsymbol{C}(A^{\mathrm{T}})$; a set of vectors $S_n = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{n-r}\}$ is a basis of $\boldsymbol{N}(A)$. Therefore, the set of vectors $S = S_r \cup S_n = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_r, \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{n-r}\}$ is a basis of $\mathbf{R}^n$.

1. • **Existence**

   Because $\mathbf{b} \in \boldsymbol{C}(A)$ and $\mathbf{b} \neq \mathbf{0}$, $\mathbf{b}$ can be expressed as a linear combination of the vectors $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n$. Let

$$\mathbf{b} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{x}_0.$$

Thus, $\mathbf{x}_0$ is a solution to $A\mathbf{x} = \mathbf{0}$. Because $\mathbf{x}_0$ is a $n-$dimensional vector, it can be expressed as a linear combination of the vectors in S. Let

$$\mathbf{x}_0 = a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_r\mathbf{r}_r + b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_{n-r}\mathbf{x}_{n-r}.$$

Let

$$\mathbf{x}_n = b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_{n-r}\mathbf{x}_{n-r}.$$

Because

$$A\mathbf{x}_n = A(b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_{n-r}\mathbf{x}_{n-r}) = b_1A\mathbf{x}_1 + b_2A\mathbf{x}_2 + \cdots + b_{n-r}A\mathbf{x}_{n-r} = \mathbf{0},$$

and

$$A\mathbf{x}_0 = A(a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_r\mathbf{r}_r + \mathbf{x}_n) = A(a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_r\mathbf{r}_r) + A\mathbf{x}_n = \mathbf{b},$$

it is clear that

$$a_1^2 + a_2^2 + \cdots + a_r^2 \neq 0.$$

Let

$$\mathbf{x}_r = \mathbf{x}_0 - \mathbf{x}_n = a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_r\mathbf{r}_r.$$

Therefore,

$$A\mathbf{x}_r = A(\mathbf{x}_0 - \mathbf{x}_n) = A\mathbf{x}_0 - A\mathbf{x}_n = \mathbf{b} - \mathbf{0} = \mathbf{b}.$$

In addition, $\mathbf{x}_r$ can be expressed as a linear combination of vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_r$, so $\mathbf{x}_r \in \boldsymbol{C}(A^{\mathrm{T}})$.

- **Uniqueness**

  Assume that there are two vectors $\mathbf{x}_{r_1}$ and $\mathbf{x}_{r_2}$ such that $\mathbf{x}_{r_1} \in \boldsymbol{C}(A^{\mathrm{T}}), \mathbf{x}_{r_2} \in \boldsymbol{C}(A^{\mathrm{T}})$ and

  $$A\mathbf{x}_{r_1} = A\mathbf{x}_{r_2} = \mathbf{b}.$$

  Therefore,

  $$A\mathbf{x}_{r_1} - A\mathbf{x}_{r_2} = A(\mathbf{x}_{r_1} - \mathbf{x}_{r_2}) = \mathbf{0},$$

  which means that $\mathbf{x}_{r_1} - \mathbf{x}_{r_2} \in \boldsymbol{N}(A)$. In addition, it is clear that $\mathbf{x}_{r_1} - \mathbf{x}_{r_2} \in \boldsymbol{C}(A^{\mathrm{T}})$. However, $\boldsymbol{N}(A)$ is the orthogonal complement of $\boldsymbol{C}(A^{\mathrm{T}})$, so we can get

  $$\mathbf{x}_{r_1} - \mathbf{x}_{r_2} = \mathbf{0} \qquad \Rightarrow \qquad \mathbf{x}_{r_1} = \mathbf{x}_{r_2}.$$

■

2. Let

$$\mathbf{x}_n = \mathbf{x}_c - \mathbf{x}_r.$$

Because $A\mathbf{x}_c = \mathbf{b}$ and $A\mathbf{x}_r = \mathbf{b}$, we have

$$A\mathbf{x}_n = A(\mathbf{x}_c - \mathbf{x}_r) = A\mathbf{x}_c - A\mathbf{x}_r = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Therefore, $\mathbf{x}_n \in \boldsymbol{N}(A)$.

■

Theorem A.15 (Property of $\boldsymbol{A^{\mathrm{T}}A}$ and $\boldsymbol{AA^{\mathrm{T}}}$) $\forall$ matrix A,

1. $A^{\mathrm{T}}A$ and AA^{T} are square and symmetric.
2. $A^{\mathrm{T}}A$ is invertible if and only if A has linearly independent columns.
3. AA^{T} is invertible if and only if A has linearly independent rows.

Proof. Assume that A is a $m \times n$ matrix, where $m, n \in \mathbb{N}^+$,

$$A = \begin{bmatrix} — & \mathbf{r}_1^{\mathrm{T}} & — \\ — & \mathbf{r}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{r}_m^{\mathrm{T}} & — \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix}.$$

1. $A^{\mathrm{T}}A$ is a $n \times n$ matrix and AA^{T} is a $m \times m$ matrix, so they are square.

   - $\forall\ 1 \le i, j \le n$ and $i, j \in \mathbb{N}^+$, the component $(A^{\mathrm{T}}A)_{ij}$ at the i-th row and j-th column of $A^{\mathrm{T}}A$ is $(A^{\mathrm{T}}A)_{ij} = \mathbf{c}_i^{\mathrm{T}}\mathbf{c}_j$; the component $(A^{\mathrm{T}}A)_{ji}$ at the j-th row and i-th column of $A^{\mathrm{T}}A$ is $(A^{\mathrm{T}}A)_{ji} = \mathbf{c}_j^{\mathrm{T}}\mathbf{c}_i$. Thus,

   $$(A^{\mathrm{T}}A)_{ij} = \mathbf{c}_i^{\mathrm{T}}\mathbf{c}_j = \mathbf{c}_j^{\mathrm{T}}\mathbf{c}_i = (A^{\mathrm{T}}A)_{ji}.$$

   - $\forall\ 1 \le i, j \le m$ and $i, j \in \mathbb{N}^+$, the component $(AA^{\mathrm{T}})_{ij}$ at the i-th row and j-th column of AA^{T} is $(AA^{\mathrm{T}})_{ij} = \mathbf{r}_i^{\mathrm{T}}\mathbf{r}_j$; the component $(AA^{\mathrm{T}})_{ji}$ at the j-th row and i-th column of AA^{T} is $(AA^{\mathrm{T}})_{ji} = \mathbf{r}_j^{\mathrm{T}}\mathbf{r}_i$. Thus,

   $$(AA^{\mathrm{T}})_{ij} = \mathbf{r}_i^{\mathrm{T}}\mathbf{r}_j = \mathbf{r}_j^{\mathrm{T}}\mathbf{r}_i = (AA^{\mathrm{T}})_{ji}.$$

   To sum up, they are symmetric.

■

2. **The proof that $\boldsymbol{N}(A^{\mathrm{T}}A) = \boldsymbol{N}(A)$.**

   - $\forall$ $n-$dimensional vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$, which means that $\mathbf{x} \in \boldsymbol{N}(A)$. Then
   $$A^{\mathrm{T}}A\mathbf{x} = A^{\mathrm{T}}\mathbf{0} = \mathbf{0},$$
   which means that $\mathbf{x} \in \boldsymbol{N}(A^{\mathrm{T}}A)$.
   - $\forall$ $n-$dimensional vector $\mathbf{x}$ such that $A^{\mathrm{T}}A\mathbf{x} = \mathbf{0}$, which means that $\mathbf{x} \in \boldsymbol{N}(A^{\mathrm{T}}A)$. Then
   $$x^{\mathrm{T}}A^{\mathrm{T}}A\mathbf{x} = (A\mathbf{x})^{\mathrm{T}}A\mathbf{x} = x^{\mathrm{T}}\mathbf{0} = \mathbf{0},$$
   which means that $A\mathbf{x} = \mathbf{0}$. Thus, $\mathbf{x} \in \boldsymbol{N}(A)$.

   To sum up, $\boldsymbol{N}(A^{\mathrm{T}}A) = \boldsymbol{N}(A)$.

   - If $A^{\mathrm{T}}A$ is invertible, then $\boldsymbol{N}(A^{\mathrm{T}}A)$ can only have $\mathbf{0}$, which is the same as $\boldsymbol{N}(A)$. Therefore, the dimension of $\boldsymbol{C}(A)$ will be n, which means that A has linearly independent columns.
   - If A has linearly independent columns, then $\boldsymbol{N}(A)$ can only have $\mathbf{0}$, which is the same as $\boldsymbol{N}(A^{\mathrm{T}}A)$. Therefore, the dimension of $\boldsymbol{C}(A^{\mathrm{T}}A)$ will be n, which means that the rank of $A^{\mathrm{T}}A$ is n and $A^{\mathrm{T}}A$ is invertible.

■

3. Assume that the matrix $B = A^{\mathrm{T}}$, so $B^{\mathrm{T}}B = AA^{\mathrm{T}}$ is invertible if and only if B has linearly independent columns. Therefore, AA^{T} is invertible if and only if A has linearly independent rows.

■

Theorem A.16 (Existence and Uniqueness of Projection) $\forall\, m, n \in \mathbb{N}^+$ such that $m \geq n$, assume that vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n \in \mathbf{R}^m$ and they are a basis of a subspace $\boldsymbol{V}$. $\forall$ vector $\mathbf{b} \in \mathbf{R}^m$ but $\mathbf{b} \notin \boldsymbol{V}$. $\exists$ a vector $\mathbf{p} \in \boldsymbol{V}$ and a vector $\mathbf{e} \in \boldsymbol{V}^{\perp}$ such that

$$\mathbf{p} + \mathbf{e} = \mathbf{b}.$$

In addition, $\mathbf{p}$ and $\mathbf{e}$ are unique.

Proof.

1. If $m = n$, then $\boldsymbol{V} = \mathbf{R}^m$, so $\mathbf{p} = \mathbf{b}$ and $\mathbf{e} = \mathbf{0}$.

2. If $m > n$, then let

$$A = \begin{bmatrix} — & \mathbf{a}_1^{\mathrm{T}} & — \\ — & \mathbf{a}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{a}_n^{\mathrm{T}} & — \end{bmatrix}.$$

   Therefore, $\mathbf{p} \in \boldsymbol{C}(A^{\mathrm{T}}) = \boldsymbol{V}$ and $\mathbf{e} \in \boldsymbol{N}(A) = \boldsymbol{V}^{\perp}$. Assume that vectors $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_{m-n}$ are a basis of $\boldsymbol{N}(A)$. Thus, the vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n, \mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_{m-n}$ form a basis of $\mathbf{R}^m$. Then $\mathbf{b}$ can be expressed as a linear combination of them. Let

$$\mathbf{b} = a_1\mathbf{a}_1 + a_2\mathbf{a}_2 + \cdots + a_n\mathbf{a}_n + b_1\mathbf{b}_1 + b_2\mathbf{b}_2 + \cdots + b_{n-m}\mathbf{b}_{n-m}.$$

   - **Existence**

     Let

$$\begin{aligned} \mathbf{p} &= a_1\mathbf{a}_1 + a_2\mathbf{a}_2 + \cdots + a_n\mathbf{a}_n, \\ \mathbf{e} &= b_1\mathbf{b}_1 + b_2\mathbf{b}_2 + \cdots + b_{n-m}\mathbf{b}_{n-m}. \end{aligned}$$

     There is $\mathbf{b} = \mathbf{p} + \mathbf{e}$.

   - **Uniqueness**

     Assume that there is another $\mathbf{p}' \in \boldsymbol{C}(A^{\mathrm{T}}) = \boldsymbol{V}$ and $\mathbf{e}' \in \boldsymbol{N}(A) = \boldsymbol{V}^{\perp}$ such that $\mathbf{b} = \mathbf{p}' + \mathbf{e}'$. Let

$$\begin{aligned} \mathbf{p}' &= a_1'\mathbf{a}_1 + a_2'\mathbf{a}_2 + \cdots + a_n'\mathbf{a}_n, \\ \mathbf{e}' &= b_1'\mathbf{b}_1 + b_2'\mathbf{b}_2 + \cdots + b_{n-m}'\mathbf{b}_{n-m}. \end{aligned}$$

     Then we have

$$\mathbf{b} - \mathbf{b} = (\mathbf{p} + \mathbf{e}) - (\mathbf{p}' + \mathbf{e}') = (\mathbf{p} - \mathbf{p}') + (\mathbf{e} - \mathbf{e}') = \mathbf{0}.$$

     There is

$$(a_1 - a_1')\mathbf{a}_1 + \cdots + (a_n - a_n')\mathbf{a}_n + (b_1 - b_1')\mathbf{b}_1 + \cdots + (b_{n-m} - b_{n-m}')\mathbf{b}_{n-m} = \mathbf{0}.$$

Because the vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n, \mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_{m-n}$ are linearly independent, there is

$$a_1 = a'_1, \quad \cdots, \quad a_n = a'_n, \qquad b_1 = b'_1, \quad \cdots, \quad b_{n-m} = b'_{n-m},$$

which means that

$$\mathbf{p} = \mathbf{p}', \qquad \mathbf{e} = \mathbf{e}'.$$

■

Theorem A.17 (Properties of Orthogonal Matrices)

1. $\forall$ orthogonal matrix Q satisfies $Q^{\mathrm{T}}Q = I$.
2. $\forall$ square orthogonal matrix Q satisfies $Q^{\mathrm{T}} = Q^{-1}$.
3. $\forall$ $m \times n$ orthogonal matrix Q, where $m, n \in \mathbb{N}^+$
   (1) $\forall$ $n-$dimensional vector $\mathbf{x}$, there is $\|Q\mathbf{x}\| = \|\mathbf{x}\|$.
   (2) $\forall$ $n-$dimensional vectors $\mathbf{x}$ and $\mathbf{y}$, there is $(Q\mathbf{x})^{\mathrm{T}}Q\mathbf{y} = \mathbf{x}^{\mathrm{T}}\mathbf{y}$.

Proof.

1. Assume that Q is a $m \times n$ orthogonal matrix, where $m, n \in \mathbb{N}^+$. Let

$$Q = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & \cdots & | \end{bmatrix}.$$

   Then

$$Q^{\mathrm{T}}Q = \begin{bmatrix} — & \mathbf{q}_1^{\mathrm{T}} & — \\ — & \mathbf{q}_2^{\mathrm{T}} & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{q}_n^{\mathrm{T}} & — \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I.$$

   ■

2. Because Q is a square orthogonal matrix, there is

$$Q^{\mathrm{T}}Q \qquad \Rightarrow \qquad Q^{\mathrm{T}} = Q^{-1}.$$

■

3. (1) There is

$$(Q\mathbf{x})^{\mathrm{T}}Q\mathbf{x} = \mathbf{x}^{\mathrm{T}}Q^{\mathrm{T}}Q\mathbf{x} = \mathbf{x}^{\mathrm{T}}(Q^{\mathrm{T}}Q)\mathbf{x} = \mathbf{x}^{\mathrm{T}}I\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{x}.$$

Therefore,

$$\left.\begin{aligned} \|Q\mathbf{x}\|^2 = (Q\mathbf{x})^{\mathrm{T}}Q\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{x} = \|\mathbf{x}\|^2 \\ \|Q\mathbf{x}\| \geq 0, \quad \|\mathbf{x}\| \geq 0 \end{aligned}\right\} \Rightarrow \qquad \|Q\mathbf{x}\| = \|\mathbf{x}\|.$$

■

(2) There is

$$(Q\mathbf{x})^{\mathrm{T}}Q\mathbf{y} = \mathbf{x}^{\mathrm{T}}Q^{\mathrm{T}}Q\mathbf{y} = \mathbf{x}^{\mathrm{T}}(Q^{\mathrm{T}}Q)\mathbf{x} = \mathbf{x}^{\mathrm{T}}I\mathbf{y} = \mathbf{x}^{\mathrm{T}}\mathbf{y}.$$

■

Theorem A.18 (Vectors in A Subspace with an Orthonormal Basis) $\forall\, n \in \mathbb{N}^+$, assume that orthonormal vectors $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n$ are a basis of a subspace $\boldsymbol{V}$. $\forall$ vector $\mathbf{v} \in \boldsymbol{V}$, there is

$$\mathbf{v} = (\mathbf{q}_1^{\mathrm{T}}\mathbf{v})\mathbf{q}_1 + (\mathbf{q}_2^{\mathrm{T}}\mathbf{v})\mathbf{q}_2 + \cdots + (\mathbf{q}_n^{\mathrm{T}}\mathbf{v})\mathbf{q}_n.$$

Proof. Because $\mathbf{v} \in \boldsymbol{V}$, $\mathbf{v}$ can be expressed as a linear combination of $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n$. Assume that

$$\mathbf{v} = k_1\mathbf{q}_1 + k_2\mathbf{q}_2 + \cdots + k_n\mathbf{q}_n.$$

$\forall\, 1 \leq i \leq n$ and $i \in \mathbb{N}^+$, there is

$$\begin{aligned} \mathbf{q}_i^{\mathrm{T}}\mathbf{v} &= \mathbf{q}_i^{\mathrm{T}}(k_1\mathbf{q}_1 + \cdots + k_{i-1}\mathbf{q}_{i-1} + k_i\mathbf{q}_i + k_{i+1}\mathbf{q}_{i+1} + \cdots + k_n\mathbf{q}_n) \\ &= k_1\mathbf{q}_i^{\mathrm{T}}\mathbf{q}_1 + \cdots + k_{i-1}\mathbf{q}_i^{\mathrm{T}}\mathbf{q}_{i-1} + k_i\mathbf{q}_i^{\mathrm{T}}\mathbf{q}_i + k_{i+1}\mathbf{q}_i^{\mathrm{T}}\mathbf{q}_{i+1} + \cdots + k_n\mathbf{q}_i^{\mathrm{T}}\mathbf{q}_n \\ &= k_i\mathbf{q}_i^{\mathrm{T}}\mathbf{q}_i \\ &= k_i\|\mathbf{q}_i\|^2 \\ &= k_i. \end{aligned}$$

Therefore,

$$k_1 = \mathbf{q}_1^{\mathrm{T}}\mathbf{v}, \qquad k_2 = \mathbf{q}_2^{\mathrm{T}}\mathbf{v}, \qquad \cdots, \qquad k_n = \mathbf{q}_n^{\mathrm{T}}\mathbf{v},$$

and

$$\mathbf{v} = (\mathbf{q}_1^{\mathrm{T}}\mathbf{v})\mathbf{q}_1 + (\mathbf{q}_2^{\mathrm{T}}\mathbf{v})\mathbf{q}_2 + \cdots + (\mathbf{q}_n^{\mathrm{T}}\mathbf{v})\mathbf{q}_n.$$

■

Theorem A.19 (The Factorization $A = QR$) $\forall\ m, n \in \mathbb{N}^+$ such that $m \geq n$, assume that vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n \in \mathbf{R}^m$ are a basis of a subspace $\boldsymbol{V}$. Gram-Schmidt constructs orthonormal vectors $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n$. The matrices with these columns satisfy $A = QR$. Then $R = Q^{\mathrm{T}}A$ is upper triangular because later $\mathbf{q}$'s are orthogonal to earlier $\mathbf{a}$'s.

$$A = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & \cdots & | \end{bmatrix}, \quad Q = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & \cdots & | \end{bmatrix}, \quad R = \begin{bmatrix} \mathbf{q}_1^{\mathrm{T}}\mathbf{a}_1 & \mathbf{q}_1^{\mathrm{T}}\mathbf{a}_2 & \cdots & \mathbf{q}_1^{\mathrm{T}}\mathbf{a}_n \\ 0 & \mathbf{q}_2^{\mathrm{T}}\mathbf{a}_2 & \cdots & \mathbf{q}_2^{\mathrm{T}}\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{q}_n^{\mathrm{T}}\mathbf{a}_n \end{bmatrix}.$$

Thus,

$$A = QR.$$

Proof. $\forall\ 1 \leq i \leq n$, where $i \in \mathbb{N}^+$, the i-th column of QR is

$$(\mathbf{q}_1^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_1 + (\mathbf{q}_2^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_2 + \cdots + (\mathbf{q}_{i-1}^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_{i-1} + (\mathbf{q}_i^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_i.$$

Then, we have

$$\mathbf{p}_i = (\mathbf{q}_1^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_1 + (\mathbf{q}_2^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_2 + \cdots + (\mathbf{q}_{i-1}^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_{i-1}.$$

Because

$$\mathbf{e}_i = \mathbf{a}_i - \mathbf{p}_i, \qquad \mathbf{q}_i = \frac{\mathbf{e}_i}{\|\mathbf{e}_i\|},$$

and $\mathbf{p}_i$ is the projection of $\mathbf{a}_i$ onto the subspace spanned by vectors $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_{i-1}$, $\mathbf{e}_i$ is the projection of $\mathbf{a}_i$ onto the line spanned by $\mathbf{q}_i$. Therefore,

$$\mathbf{e}_i = (\mathbf{q}_i^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_i.$$

Thus,

$$\mathbf{a}_i = (\mathbf{q}_1^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_1 + (\mathbf{q}_2^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_2 + \cdots + (\mathbf{q}_{i-1}^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_{i-1} + (\mathbf{q}_i^{\mathrm{T}}\mathbf{a}_i)\mathbf{q}_i.$$

Now we have

$$A = QR.$$

■