ECE 707: Linear Systems (1)

T. Kirubarajan

Department of Electrical and Computer Engineering
McMaster University
Hamilton, Ontario, Canada

These viewgraphs are based on the text
"Linear System: Theory and Design" by Chi-Tsong Chen
Oxford University Press, 1999.

1. Introduction

System Theory: Deals with mathematical representation of input-output relationships (or internal representation).

Model: Describes the evolution of the system over time and I/O relationship at any time using vectors, matrices and functions.

Used for

interpretation of past behavior (estimation)
prediction of future behavior(prediction)
modification of behavior (control)
accumulation of knowledge (learning).

Purpose: Understand behavior and eventually improve it.

Rectangular array $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]$ is $m \times n$.

a_{ij} element in row i and column j. $a_{ij} \in$ field, F.

Field: A set of scalars that include an identity element “1” and a zero element “0” with addition and multiplication operations so that

(i) $\alpha, \beta \in F \Rightarrow \alpha + \beta \in F,\ \alpha \cdot \beta \in F$

(ii) $\alpha + 0 = \alpha,\ \alpha \cdot 0 = 0,\ \alpha \cdot 1 = \alpha\ \ \forall \alpha \in F$.

There exists an unique negative $(-\alpha)$ and unique inverse (α^{-1}) for each $\alpha \in F$. such that

(iii) $\alpha + (-\alpha) = 0$

(iv) $\alpha \cdot (\alpha^{-1}) = (\alpha^{-1}) \cdot \alpha = 1$ for $\alpha \neq 0$.

Finally, commutative, associative, distributive properties of algebra holds for $+, \cdot$ operators.

Elements of matrices $A,\ B,\ C$ are from a field F

(i) $A+B=[a_{ij}+b_{ij}]=[b_{ij}+a_{ij}]=B+A$ ("+" operation commutative)

(ii) $A+(B+C)=(A+B)+C$ ("+" operation associative)

(iii) $A\cdot(B+C)=AB+AC$ (distributive)

(iv) $\alpha\cdot(A+B)=\alpha A+\alpha B$ (distributive) also $(\alpha+\beta)\cdot A=\alpha A+\beta A$

(v) $\alpha\cdot(\beta\cdot A)=(\alpha\cdot\beta)\cdot A$ (associative)

But $AB\neq BA$ in general, i.e., "." operation is not commutative for matrices.

Some Special Operations on Matrices

Transpose of a matrix $A=A^T$: $A=[a_{ij}],\ A^T=[a_{ji}]$.

Conjugation of a matrix $A=A^*$: $A=[a_{ij}],\ A^*=[a_{ji}^*]$.

Trace of a matrix $A=\text{Tr}(A)=\sum_{i=1}^{n}a_{ii}$ (only for square matrices).

$\frac{dA(t)}{dt}=\left[\frac{da_{ij}(t)}{dt}\right]$ and similarly $\int_{t_1}^{t_2}A(t)dt=\left[\int_{t_1}^{t_2}a_{ij}(t)dt\right]$.

Special Matrices

(1) Identity $I_n = n \times n$ matrix $[\delta_{ij}]$; where $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & \text{else} \end{cases}$

(2) Diagonal $D = n \times n$ matrix $d_{ij}]$ where $d_{ij} = \begin{cases} d_i, & i = j \\ 0, & \text{else} \end{cases}$

(3) Vector $\underline{x} = n \times 1$ matrix (column matrix).

(4) Real matrix $A = A^* \Rightarrow a_{ij} = a_{ij}^*$.

(5) Symmetric $A = A^T \Rightarrow a_{ij} = a_{ji}$.

(6) Hermitian $A = A^{*T} = A^H \Rightarrow a_{ij} = a_{ji}^*$.

Some Useful Properties

(1) $\underline{x}^T \underline{y} = \underline{y}^T \underline{x}$ (2) $\underline{x}^{*T} \underline{x} = 0 \Leftrightarrow \underline{x} = 0$ (3) $(AB)^T = B^T A^T$.

McMaster University

LVS over a field F denoted as (X, F) consists of a set X of vectors, a field F and 2 operations vector addition and scalar multiplication that satisfy

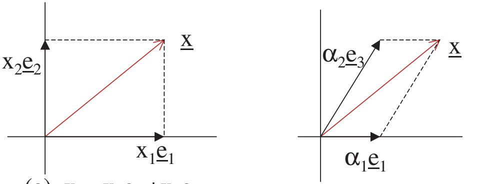
1. $\forall \underline{x}_1, \underline{x}_2 \in X \quad \underline{x}_1 + \underline{x}_2 \in X$ (closure under vector addition)
2. $\underline{x}_1 + \underline{x}_2 = \underline{x}_2 + \underline{x}_1$ (addition is commutative)
3. $(\underline{x}_1 + \underline{x}_2) + \underline{x}_3 = \underline{x}_1 + (\underline{x}_2 + \underline{x}_3)$ (addition is associative)
4. $\exists \underline{0} \in X$ such that $\underline{0} + \underline{x} = \underline{x} \ \ \forall \underline{x} \in X$ (zero vector)
5. $\forall \underline{x} \in X \ \ \exists (-\underline{x}) \in X$ such that $\underline{x} + (-\underline{x}) = \underline{0}$ (negative of a vector)
6. $\forall \alpha \in F$ and $\forall \underline{x} \in X \ \ \alpha \underline{x} \in X$
7. $\forall \alpha, \beta \in F$ and $\forall \underline{x} \in X \ \ \alpha(\beta \underline{x}) = (\alpha\beta)\underline{x}$ (scalar multiplication is associative)
8. $\alpha(\underline{x}_1 + \underline{x}_2) = \alpha \underline{x}_1 + \alpha \underline{x}_2$ (scalar multiplication is distributive w.r.t vector addition)
9. $\forall \underline{x} \in X \ \ 1\underline{x} = \underline{x}$ ("1" is the unity element in F)

Example $(\Im^n, \Re)$ or $(\Im^n, \Im)$ but not $(\Re^n, \Im)$ why?

Linear Dependence: The set of vectors $\{\underline{x}_i\}_{i=1}^n$ in the LVS (X, F) is linearly dependent if $\exists\{\alpha_i\}_{i=1}^n$, not all zero, such that $\sum \alpha_i \underline{x}_i = \underline{0}$.
If the above holds only for $\alpha_i = 0 \;\; \forall i$ then $\{\underline{x}_i\}_{i=1}^n$ are linearly independent. This can be expressed as

$$\underline{\alpha} \stackrel{\triangle}{=} \text{col}(\alpha_1, \cdots, \alpha_n), \quad [\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n]\underline{\alpha} = \underline{0} \Leftrightarrow \underline{\alpha} = \underline{0} \tag{1}$$

Dimension of the LVS: Maximum number of linearly independent vectors in the space.
For the LVS $(\Re^2, \Re)$ any vector $\underline{x}$ can be represented by different sets of linearly independent vectors as shown in Fig. 1.1.



(a) $\underline{x} = x_1\underline{e}_1 + x_2\underline{e}_2$ (b) $\underline{x} = \alpha_1\underline{e}_1 + \alpha_2\underline{e}_3$

Example: Consider a linear vector space $(\Re^n, \Re)$.

Let $\underline{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, the element in row i is 1 and rest are zeros, where $i = 1, \cdots, n$.

Here $\sum_{i=1}^{n} \alpha_i \underline{e}_i = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \underline{0} \Rightarrow \underline{\alpha} = \underline{0}$. Thus $\{\underline{e}_i\}_{i=1}^{n}$ are linearly independent.

For any vector $\underline{x} \in \Re^n$, $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i \underline{e}_i$.

Hence $\{\underline{x}, \underline{e}_1, \underline{e}_2, \cdots, \underline{e}_n\}$ are not linearly independent for any $\underline{x}$.
Dimension of LVS $(\Re^n, \Re)$ is n.

Span of a LVS: Vectors $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n$ span a LVS if every $\underline{x} \in X$ can be expressed as a linear combination $\underline{x} = \sum_{i=1}^{n} \alpha_i \underline{x}_i$, for $\alpha_i \in F$.
Example: $\underline{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \underline{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \underline{x}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
Since $\underline{x}_1 + \underline{x}_2 - \underline{x}_3 = \underline{0} \Rightarrow$ vectors are not Linearly independent.
For any $\underline{y} \in \Re^2$

$$\underline{y} = \begin{bmatrix} a \\ b \end{bmatrix} = \alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \alpha_3 \underline{x}_3 = \begin{bmatrix} \alpha_1 + \alpha_3 \\ \alpha_2 + \alpha_3 \end{bmatrix} \quad \forall a, b$$

provided $\alpha_1 = a - \alpha_3,\ \alpha_2 = b - \alpha_3$. Thus $\{\underline{x}_1, \underline{x}_2, \underline{x}_3\}$ span the LVS $(\Re^2, \Re)$. However, the choice of α_i's is not unique.

Basis: Any set of k lin. indep. vectors $\{\underline{x}_1, \underline{x}_2 \cdots, \underline{x}_n\}$ of a LVS of dimension k is called a basis for the space.
Example: $\{\underline{e}_1, \underline{e}_2, \cdots, \underline{e}_n\}$ forms a basis of $(\Re^n, \Re)$.

Theorem: Any vector in a LVS can be expressed as an unique linear combination of the basis vectors.

Proof: Let $\{\underline{x}_1, \underline{x}_2 \cdots, \underline{x}_k\}$ be a basis of (X, F). Then for any $\underline{y} \in X$

$$\alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \cdots + \alpha_k \underline{x}_k + \alpha_0 \underline{y} = \underline{0} \text{ for some } \underline{\alpha} \neq 0$$

Otherwise dimension $> k$. Also $\alpha_0 \neq 0$; otherwise basis vectors are not linearly independent. Hence

$$\underline{y} = -\frac{1}{\alpha_0} \sum_{i=1}^{k} \alpha_i \underline{x}_i = \sum_{i=1}^{k} \beta_i \underline{x}_i$$

Suppose β_i are not unique, i.e., $\underline{y} = \sum_{i=1}^{k} \beta_i \underline{x}_i = \sum_{i=1}^{k} \gamma_i \underline{x}_i,\ \underline{\beta} \neq \underline{\gamma}$. Then $\sum_{i=1}^{n} (\beta_i - \gamma_i) \underline{x}_i = \underline{0}$ or, $\{\underline{x}_1, \underline{x}_2 \cdots, \underline{x}_k\}$ are not lin. indep., a contradiction.

Thus "representation" $\underline{\beta}$ of $\underline{y}$ for the basis $\{\underline{x}_1, \underline{x}_2 \cdots, \underline{x}_k\}$ is unique.

Example: Let
$X =$ set of polynomials of degree $n-1$ or less in variable s and
$F = \Re$.
Let $e_i = s^{i-1},\ i = 1, 2, \cdots, n$. Prove that e_i's form a basis for $(X, \Re)$.
Proof: For an arbitrary $x \in X$

$$x = \sum_{i=1}^{n} \alpha_i e_i = \sum_{i=1}^{n} \alpha_i s^{i-1} = [e_1\ e_2 \cdots e_n]\underline{\alpha}$$

Clearly e_i's span $(X, \Re)$. To prove that they form a basis, we need to further prove that they are lin. indep.
Lin. indep. of $\{e_1 \cdots e_n\}$ requires $x = 0$ for all s has unique solution $\underline{\alpha} = \underline{0}$.
As $|s| \to \infty$, $x \to \alpha_n e_n = \alpha_n s^{n-1}$, or, $x = 0$ means $\alpha_n = 0$.
For $\alpha_n = 0$, $x \to \alpha_{n-1} s^{n-2}$ when $|s| \to \infty$, i.e., $x = 0 \Rightarrow \alpha_{n-1} = 0$, so on.
Thus $x = 0 \Rightarrow \underline{\alpha} = \underline{0}$. (proved)

Let $T(\ )$ be a transformation of $\underline{x} \in X$ to $\underline{y} \in Y$; i.e., $\underline{y} = T(\underline{x})$. $\underline{x}$ is called "domain" of T and Y is called "range space" of T.

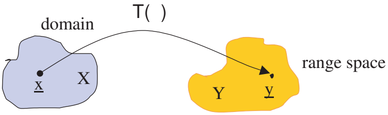


Fig. 1.2

The transformation $T(\ )$ is linear if and only if

(i) $T(\underline{x}_1 + \underline{x}_2) = T(\underline{x}_1) + T(\underline{x}_2)\ \ \forall \underline{x}_1, \underline{x}_2 \in X$

(ii) $T(\alpha \underline{x}) = \alpha T(\underline{x})\ \ \forall \underline{x} \in X, \alpha \in F$.

$T(\ )$ linear and (X, F) a LVS $\Rightarrow$ (Y, F) a LVS. (Hint: Prove the properties of LVS for $\forall \underline{y} \in Y$)

In finite dimension $T(\ )$ can be expressed by a matrix.

Proof: Without loss of generality (WLOG) we consider

$$\underline{x} = [\underline{u}_1 \cdots \underline{u}_n]\underline{\alpha} = U\underline{\alpha} \ \ \forall \underline{x} \in \Im^n, \alpha_i \in \Im$$
$$\underline{y} = [\underline{w}_1 \cdots \underline{w}_m]\underline{\beta} = W\underline{\beta} \ \ \forall \underline{y} \in \Im^m, \beta_i \in \Im$$

Then $T(\underline{x}) = T(\sum_{i=1}^n \alpha_i \underline{u}_i) = \sum_{i=1}^n \alpha_i T(\underline{u}_i)$. (note that $\underline{u}_i$'s form a basis of $\Im^n$.)

Let $T(\underline{u}_i) = W\underline{b}_i, \ i = 1, 2, \cdots, n$. (Since columns of W forms a basis for $\Im^m$.)

We have $W\underline{\beta} = T(\underline{x}) = \sum_{i=1}^n \alpha_i W\underline{b}_i = W[\underline{b}_1 \ \underline{b}_2 \cdots \underline{b}_n]\underline{\alpha} \stackrel{\triangle}{=} WB\underline{\alpha}$
This means $\underline{\beta} = B\underline{\alpha}$. (Since W is invertible. Special case $W = [\underline{e}_1 \cdots \underline{e}_m] = I_m$.)
Therefore $m \times n$ matrix $B = [\underline{b}_1 \ \underline{b}_2 \cdots \underline{b}_n]$ is a linear operator that maps say $(\Im^n, \Im)$ to $(\Im^m, \Im)$.

Symbol det(A) or $|A|$, only defined for square matrices.
If A is $n \times n$, then det(A) is n-fold product of the elements of A.
Each term of the product contains one and only one element from each row and each column of A.
Sign of each term depends the sequence of on leading and trailing subscripts.

Example: $$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}$$

First set second (trailing) subscript in ascending order ($1 \rightarrow 2 \rightarrow 3$).
Now the sign depends on the number of reversals of the leading subscripts from the order ($1 \rightarrow 2 \rightarrow 3 \rightarrow 1$). For even number of reversals sign is "$+$" and for odd number of reversals sign is "$-$".
Example: $a_{21}a_{12}a_{33}$ has $2 \rightarrow 1$, $1 \rightarrow 3$, $3 \rightarrow 2$. 3 reversals, sign -1.

Can prove $|AB| = |A| \cdot |B|$. The tedious proof is avoided but the result is important.

Example: Consider Any matrix $A_{n\times n}$ and a special matrix $E_{n\times n}$ defined below.

$$\underbrace{\begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}}_{E} \cdot \underbrace{\begin{bmatrix} \text{row } 1 \\ \text{row } 2 \\ \vdots \\ \text{row } n \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \sum_{i=1}^{n} \alpha_{1i}\text{row } i \\ \text{row } 2 \\ \vdots \\ \text{row } n \end{bmatrix}}_{EA}$$

We have $|EA| = |E||A| = \alpha_{11}|A|$.

From the above we have following properties of determinant

(1) Replacing row j by $\sum_{i=1}^{n} \alpha_{ji} \times$ row i changes $|A|$ by α_{jj}.

(2) If rows (or columns) of A are not linearly independent then $|A| = 0$.

Proof: (1) is a general case of the example. (2) can be proved by choosing β_i such that $\sum_{i=1}^{n} \beta_i \times \text{row}_i = \underline{0}^T$.

Adjoint of a Matrix

Let a $n \times n$ matrix $A = [a_{ij}]$.

$M_{ij} \stackrel{\triangle}{=}$ determinant of A less row i and column j (i, j th minor of A).

$\Delta_{ij} \stackrel{\triangle}{=} (-1)^{i+j} M_{ij}$ (cofactor i, j of A).

Adjoint of A is defined as

$$\text{adj}(A) = \begin{bmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & \ldots & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{bmatrix} = [\Delta_{ji}]$$

Laplace Expansion of $|A|$

$$|A| = \sum_{i=1}^{n} a_{ij} \Delta_{ij} = \sum_{j=1}^{n} a_{ij} \Delta_{ij}$$

Matrix Inverse

Let $B = \text{adj}(A)$ and $C = AB$. Then we have

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} a_{ik} \Delta_{jk} = \begin{cases} 0, & i \neq j \\ |A|, & i = j \end{cases}$$

$i = j$ part follows from Laplace expansion over row i.
When $i \neq j$ the expression is equal to det($\tilde{A}$) where j th row of A is replaced by its i th row. As a result the rows become linearly dependent.
Hence $A \cdot \text{adj}(A) = |A| I_n$.
If $|A| \neq 0$ then $A \frac{\text{adj}(A)}{|A|} = I$ or $A^{-1} = \frac{\text{adj}(A)}{|A|}$.

If A^{-1} exists $\Leftrightarrow |A| \neq 0$.

Proof: $\Leftarrow \; A^{-1} = \frac{\text{adj}(A)}{|A|}$ if $|A| \neq 0$
$\Rightarrow$ suppose A^{-1} exists then $|A||A^{-1}| = |AA^{-1}| = |I| = 1 \Rightarrow |A| \neq 0$

If $|A| \neq 0$ then $A^{-1} = \frac{\text{adj}(A)}{|A|}$ is unique.

Proof: Suppose $BA = CA = I, \; |A| \neq 0$.

Then $BA \cdot \text{adj}(A) = CA \cdot \text{adj}(A) \;\; \Rightarrow B|A| = C|A| \;\; \Rightarrow B = C$.

Rank of a Matrix $A_{m\times n}$

If $PAQ = \begin{array}{cc} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array}\right] \end{array}$ where $r \leq \min(m,n)$.

Also if $P_{m\times m}$ and $Q_{n\times n}$ are non-singular matrices then rank of A is r.

A has r linearly independent rows and columns.

Proof: We have

$$PA = \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array}\right] Q^{-1} = \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array}\right] \left[\begin{array}{cc} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{21} & \tilde{Q}_{22} \end{array}\right] = \begin{array}{cc} \left[\begin{array}{cc} \tilde{Q}_{11} & \tilde{Q}_{12} \\ 0 & 0 \end{array}\right] & \begin{array}{c} r \\ m-r \end{array} \end{array}$$

(1) Since Q^{-1} exists its rows are linearly independent.

(2) Since P is non-singular $A, \; PA$ has same number of lin. indep. rows.

The above equation and the rules prove A has r lin. indep. rows.

Summary: $A_{m\times n}$ has equal number of lin. indep. rows and columns, $r \leq \min(m,n)$.

Non-singular matrices $P,\ Q$ exist to yield $PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

Special Case: $A_{n\times n}$ or square matrix.

If $r = n$, A has n lin. indep. rows and columns and A^{-1} exists.

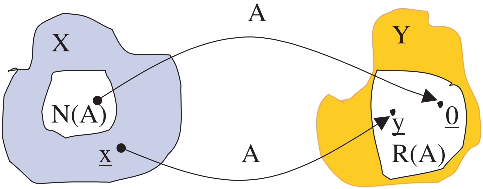


Fig. 1.3

Range Space of A

$$R(A) \stackrel{\triangle}{=} \{\underline{y} | A\underline{x} = \underline{y},\ \underline{x} \in X\}$$

For $A_{m\times n} = [a_{ij}]$ if $a_{ij} \in \Im$ and $X \in \Im^n$ then $R(A)$ is a linear subspace of $\Im^m$, i.e., a LVS contained in $(\Im^m, \Im)$.

Proof: (1) Let $\underline{y}_1, \underline{y}_2 \in R(A)$ where $A\underline{x}_i = \underline{y}_i$.

Then $\underline{y}_1 + \underline{y}_2 = A\underline{x}_1 + A\underline{x}_2 = A(\underline{x}_1 + \underline{x}_2) \in R(A)$.

So $\forall \underline{y}_1, \underline{y}_2 \in R(A),\ \underline{y}_1 + \underline{y}_2 \in R(A)$.

(2) For every $\alpha \in \Im$ and $\underline{y} \in R(A)$

$\alpha\underline{y} = \alpha A\underline{x} = A(\alpha\underline{x}) \in R(A)$.

We can also show that the commutative, associative and distributive properties of vector addition and scalar multiplication holds.
Again $A\underline{0} = \underline{0}$, hence $R(A)$ contains zero vector.
Finally if $A\underline{x} = \underline{y}$, then $A(-\underline{x}) = (-\underline{y})$. Thus for every $\underline{y} \in R(A)$, $(-\underline{y}) \in R(A)$.
The above discussion proves that $(R(A), \Im)$ forms a LVS.
Note: $\underline{y} = A\underline{x} = [\underline{a}_1\ \underline{a}_2 \cdots \underline{a}_n]\underline{x} = \sum_{i=1}^{n} x_i \underline{a}_i$, i.e., $\underline{y}$ is a linear combination of the columns of A.
Hence $R(A)$ is spanned by the columns of A.
$\rho(A) =$dimension of $R(A) =$rank of A.

Null Space of A

$$N(A) \stackrel{\triangle}{=} \{\underline{x} \in X | A\underline{x} = \underline{0}\}$$

$\gamma(A) =$dimension of null space of A.

$$A_{m\times n} \Rightarrow \rho(A) + \gamma(A) = n$$

Proof: Consider $A = [\underline{a}_1\ \underline{a}_2 \cdots \underline{a}_r\ \underline{a}_{r+1} \cdots\ \underline{a}_n];\ \rho(A) = r$.
Assume for simplicity that $\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_r$ are linearly independent and which means rest are not.
Then $\underline{a}_j = \beta_{j1}\underline{a}_1 + \cdots + \beta_{jr}\underline{a}_r = \sum_{i=1}^{r} \beta_{ji}\underline{a}_i,\ j = r+1, \cdots, n$.
or $-\sum_{i=1}^{r} \beta_{ji}\underline{a}_i + \underline{a}_j = \underline{0},\ j = r+1, \cdots, n$

$$\Rightarrow A\underline{v}_j = \underline{0},\ j = r+1, \cdots, n \text{ where } \underline{v}_j = \begin{bmatrix} -\beta_{j1} \\ \vdots \\ -\beta_{jr} \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Thus $\underline{v}_{r+1}, \underline{v}_{r+2}, \cdots, \underline{v}_n$ are in $N(A)$ and are linearly independent by construction. Hence $\gamma(A) \geq n - r$.

Consider any vector $\underline{s} \in N(A)$. Then

$$A\underline{s} = [\underline{a}_1 \cdots \underline{a}_r \ \underline{a}_{r+1} \cdots \underline{a}_n]\underline{s} = \sum_{i=1}^{r} \underline{a}_i s_i + \sum_{j=r+1}^{n} \underline{a}_j s_j = \underline{0}$$

Since $\underline{a}_j$ must satisfy $\underline{a}_j = \sum_{i=1}^{r} \beta_{ji}\underline{a}_i, \ j = r+1, \cdots, n$, we have

$$\sum_{i=1}^{r} \underline{a}_i s_i + \sum_{j=r+1}^{n} s_j \sum_{j=1}^{r} \underline{a}_i \beta_{ji} = \sum_{i=1}^{r} \underline{a}_i (s_i + \sum_{j=r+1}^{n} s_j \beta_{ji}) = \underline{0}$$

However $\underline{a}_1, \cdots, \underline{a}_r$ are lin. indep. So this requires that $s_i + \sum_{j=r+1}^{n} s_j \beta_{ji} = 0$

Thus $\underline{s} = \begin{bmatrix} -\sum_{i=r+1}^{n} s_j \beta_{j1} \\ \vdots \\ -\sum_{i=r+1}^{n} s_j \beta_{jr} \\ s_{r+1} \\ \vdots \\ s_n \end{bmatrix} = \sum_{j=r+1}^{n} \underline{v}_j s_j = [\underline{v}_{j+1} \cdots \underline{v}_n] \begin{bmatrix} s_{r+1} \\ \vdots \\ s_n \end{bmatrix}$

Indicating that any $\underline{s} \in N(A)$ is a lin. comb. of $\underline{v}_{r+1}, \cdots, \underline{v}_n$. Hence $\gamma(A) = n - r$.

Example: Find a set of basis for the null space of A.

$$A_{3\times 5} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Solution: (Clearly $\rho(A) = 3,\ n = 5,\ \gamma(A) = 5 - 3 = 2.$)

Consider $A\underline{x} = \underline{0} \ \Rightarrow \begin{cases} x_1 + x_4 = 0 \\ x_2 + x_4 + 2x_5 = 0 \\ x_3 + x_5 = 0 \end{cases}$

$\underline{x}$ can be written as, $\underline{x} = [x_1 \ \ x_1 + 2x_2 \ \ x_3 \ \ -x_1 \ \ -x_3]^T$ Clearly x_1 and x_3 can be chosen arbitrarily.

Two independent variables mean $N(A)$ has two dimensions. Basis vectors can be obtained by choosing two sets of linearly independent values of $\{x_1, x_3\}$.

For example, $\underline{y}_1 = [1\ 1\ 0\ -1\ 0]^T$ ($x_1 = 1, x_3 = 0$);

$\underline{y}_2 = [0\ 2\ 1\ 0\ -1]^T$ ($x_1 = 0, x_3 = 1$). Therefore $\{\underline{y}_1, \underline{y}_2\}$ forms a basis for $N(A)$.

If $C = AB$, then $\rho(C) \leq \min(\rho(A), \rho(B))$.

Proof: Let us consider $A_{m\times n},\ B_{n\times q}$ and $\rho(A) = r,\ \rho(B) = p$.
It can be shown that columns of C are linear combinations of columns of A.
Furthermore, columns of A span r dimensional space. So columns of C span a space whose dimension is $\leq r$. Thus $\rho(C) \leq r$.
Again it can be shown that rows of C are linear combinations of rows of B. Using similar argument we can prove that $\rho(C) \leq p$.
So we have
$\rho(C) \leq r, p \ \Rightarrow \rho(C) \leq \min(r, p) \ \Rightarrow \rho(C) \leq \min(\rho(A), \rho(B))$.

If P non-singular then $\rho(PA) = \rho(A)$.

Proof: Consider $P_{m\times m}$ and $A_{m\times n}$.
Then $\rho(PA) \leq \min(\rho(P), \rho(A)) = \min(m, \rho(A)) = \rho(A)$ as $\rho(A) \leq \min(m, n)$.
Again $\rho(A) = \rho(P^{-1}PA) \leq \min\left(\rho(P^{-1}), \rho(PA)\right) \leq \rho(PA)$.
Thus we have $\rho(PA) \leq \rho(A)$ and $\rho(PA) \geq \rho(A)$. So $\rho(PA) = \rho(A)$.

Similarly it can be shown that for non-singular Q, $\rho(AQ) = \rho(A)$.
In general, multiplication of A with non-singular matrices does not change it's rank.

Linear Systems of Equations: $A\underline{x} = \underline{y}$ **where** A **is** $m \times n$

$A\underline{x} = \underline{y}$ has a solution $\Leftrightarrow \underline{y} \in R(A)$ as when a solution exists $\underline{y}$ is linear combination of columns of A.

Let $W \stackrel{\triangle}{=} [A\ y]_{m\times n+1}$. Two cases arise

(1) If $\rho(W) < \rho(A)$, no solution exists since $\underline{y} \notin R(A)$, i.e., not in a subspace spanned by columns of A.

(2) If $\rho(W) = \rho(A)$, $\underline{y} \in R(A)$ and atleast one solution exists. Consider two cases:

(a) $\rho(A) = n$, $\gamma(A) = 0$ ($m \geq n$). Solution is unique.
Proof: Suppose $A\underline{x}_1 = A\underline{x}_2 = \underline{y}$, $\underline{x}_1 \neq \underline{x}_2$. Then $A(\underline{x}_1 - \underline{x}_2) = \underline{0}$, with $\underline{x}_1 - \underline{x}_2 \neq \underline{0}$ which violets the fact that null space is empty.

(b) $\rho(A) < n$, $\gamma(A) > 0$. Solution not unique. The difference between solutions for example $\underline{x}_1 - \underline{x}_2$ must be in null space $N(A)$.

Solution of $A\underline{x} = \underline{y}$, for $A_{n\times n}$, $|A| \neq 0$

The solution is $\underline{x} = A^{-1}\underline{y}$, where $A^{-1} = \frac{\text{adj}(A)}{|A|}$.
Direct way of finding $\underline{x}$ is computationally expansive.
It can be proved that for any $A_{n\times n}$ we can have non-singular P such that $PA = U$, U =upper triangular matrix.
Then $U\underline{x} = P\underline{y} \stackrel{\triangle}{=} \underline{z}$. Now can solve $U\underline{x} = \underline{z}$ due to nature of U. (Note that $u_{ii} \neq 0$ if $|A| \neq 0$.)

$$\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1,n-1} & u_{1n} \\ 0 & u_{22} & \cdots & u_{2,n-1} & u_{2n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & u_{n-1,n-1} & u_{n-1,n} \\ 0 & 0 & \cdots & 0 & u_{nn} \end{bmatrix} \underline{x} = \underline{z}$$

This gives $x_n = z_n/u_{nn}$; $x_{n-1} = (z_{n-1} - u_{n-1,n}z_n)/u_{n-1,n-1}$;
$\cdots\ x_1 = (z_1 - u_{12}x_2 - u_{13}x_3 - \cdots - u_{1n}x_n)/u_{11}$.
P, U are not unique. These can be found in various ways.

(a) Gauss Elimination

Step 1. Let $P_1 A \triangleq A_1 = [a^1_{ij}]$ interchange row 1, row j where $|a_{j1}| = \max_i |a_{i1}|$.

Step 2. Let $P_2 A_1 = A_2 = [a^2_{ij}]$ zero out column 1 below (1,1) element of A_1.

$$P_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -a^1_{21}/a^1_{11} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ -a^1_{n1}/a^1_{11} & 0 & 0 & \cdots & 1 \end{bmatrix}, \; P_2 A_1 = \begin{bmatrix} a^1_{11} & a^1_{12} & \cdots & a^1_{1n} \\ 0 & a^2_{22} & \cdots & a^2_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ 0 & a^2_{n2} & \cdots & a^2_{nn} \end{bmatrix} = A_2$$

Repeat this process an last $n-1$ rows of A_2, etc.

Repeat $n-2$ times until $P = P_{2(n-1)} \cdots P_2 P_1, \; PA = U$.

(b) Householder Transform

Consider $A = [\underline{a}_1 \; \underline{a}_2 \cdots \underline{a}_n]$.

Let $\underline{u}_1 = \begin{bmatrix} \sqrt{\underline{a}_1^{*T}\underline{a}_1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\underline{w} = \underline{a}_1 - \underline{u}_1, \; P_1 = I_n - \frac{2\underline{w}\,\underline{w}^{*T}}{\underline{w}^{*T}\underline{w}}$

Then $P_1 A = \begin{bmatrix} \sqrt{\underline{a}_1^{*T}\underline{a}_1} & \mathsf{x}\cdots\mathsf{x} \\ \underline{0} & A_2 \end{bmatrix}$ where $A_2 = [\underline{a}_1^2 \; \underline{a}_2^2 \cdots \underline{a}_{n-1}^2]$.

Let

$\underline{u}_2 = \begin{bmatrix} \sqrt{(\underline{a}_1^2)^{*T}\underline{a}_1^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\underline{w}_2 = \underline{a}_1^2 - \underline{u}_2, \; P_2 = \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \tilde{P}_2 \end{bmatrix}$, $\tilde{P}_2 = I_{n-1} - \frac{2\underline{w}_2\underline{w}_2^{*T}}{\underline{w}_2^{*T}\underline{w}_2}$

Then $P_2 P_1 A = \begin{bmatrix} \sqrt{\underline{a}_1^{*T}\underline{a}_1} & 0 & \mathsf{x}\cdots\mathsf{x} \\ 0 & \sqrt{(\underline{a}_1^2)^{*T}\underline{a}_1^2} & \mathsf{x}\cdots\mathsf{x} \\ \underline{0} & \underline{0} & A_3 \end{bmatrix}$

Continue until $P_{n-1}P_{n-2}\cdots P_1 A = U$.

Example: Find $\underline{x}$ using the following information.

$$A\underline{x} = \underline{y};\ A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix},\ \underline{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

(1) Gauss Elimination

$$P_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},\ P_1 A = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix},\ P_2 P_1 A = \begin{bmatrix} 2 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & 3/2 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix},\ P_3 P_2 P_1 A = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 5/2 \end{bmatrix}$$

$$P_3P_2P_1\underline{y} = P_3P_2\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = P_3\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We have $\begin{bmatrix} 2 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 5/2 \end{bmatrix}\underline{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

or, $5/2x_3 = 1 \Rightarrow x_3 = 2/5;\ x_2 - x_3 = 0 \Rightarrow x_2 = 2/5;$
$2x_1 + 2x_2 + x_3 = 0 \Rightarrow x_1 = -3/5$.

(2) Householder Transform

$$\underline{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix},\ \underline{a}_1^T\underline{a}_1 = 9,\ \underline{u} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \underline{w} = \underline{a}_1 - \underline{u} = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}$$

$$P_1 = I_3 - 2\frac{\underline{w}\,\underline{w}^T}{\underline{w}^T\underline{w}} = 1/3 \times \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix},\ P_1A = \begin{bmatrix} 3 & 8/3 & 8/3 \\ 0 & 1/3 & 4/3 \\ 0 & 4/3 & 1/3 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1/3 & 4/3 \\ 4/3 & 1/3 \end{bmatrix},\ \underline{a}_1^2 = \begin{bmatrix} 1/3 \\ 4/3 \end{bmatrix},\ \underline{u} = \begin{bmatrix} 1.3744 \\ 0 \end{bmatrix},\ \underline{w} = \underline{a}_1^2 - \underline{u} = \begin{bmatrix} -1.0410 \\ 4/3 \end{bmatrix}$$

$$\tilde{P}_2 = I_2 - 2\frac{\underline{w}\,\underline{w}^T}{\underline{w}^T\underline{w}} = \begin{bmatrix} 0.2425 & 0.9701 \\ 0.9701 & -0.2425 \end{bmatrix},\ P_2 = \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \tilde{P}_2 \end{bmatrix}$$

$$P_2 P_1 A = \begin{bmatrix} 3 & 8/3 & 8/3 \\ 0 & 1.3743 & 0.6467 \\ 0 & 0 & 1.2126 \end{bmatrix},\ P_2 P_1 \underline{y} = \begin{bmatrix} 1/3 \\ 0.8084 \\ 0.4851 \end{bmatrix}$$

Therefore $\begin{bmatrix} 3 & 8/3 & 8/3 \\ 0 & 1.3743 & 0.6467 \\ 0 & 0 & 1.2126 \end{bmatrix} \underline{x} = \begin{bmatrix} 1/3 \\ 0.8084 \\ 0.4851 \end{bmatrix} \Rightarrow \underline{x} = \begin{bmatrix} -3/5 \\ 2/5 \\ 2/5 \end{bmatrix}$