2. Linear Transformations & Matrices

Thursday, April 6, 2023 5:06 PM

Linear Transformations

Symmetries (about a line passing through the origin) and rotations (about the origin) are mappings that are "Linear."

Definition:

A function $L: R^m \rightarrow R^n$ is linear if

1. for all $v, w \in R^m$ we have $L(v+w) = L(v) + L(w)$ and
2. for all $v \in R^m$ and all $\alpha \in R$ we have $L(\alpha v) = \alpha L(v)$

Example: To show

$$L: R^2 \longrightarrow R^3 \quad \text{is linear?}$$
$$(v_1, v_2) \longrightarrow (5v_1, 0, v_1+v_2)$$

Ans Summation property:

Let there be 2 points in R^2: $x: (x_1, x_2)$, $y: (y_1, y_2)$

$$L(x+y) = L((x_1+y_1), (x_2+y_2))$$
$$= (5(x_1+y_1), 0, x_1+x_2+y_1+y_2) \text{ — (1)}$$

$$L(x) = (5x_1, 0, x_1+x_2) \text{ — (2)}$$
$$L(y) = (5y_1, 0, y_1+y_2) \text{ — (3)}$$

$$L(x) + L(y) \Rightarrow (2) + (3)$$
$$= (5(x_1+y_1), 0, x_1+y_1+x_2+y_2)$$
$$= (1)$$

Thus summation property holds!

$$\alpha x = (\alpha x_1, \alpha x_2)$$
$$\therefore L(\alpha x) = (5\alpha x_1, 0, \alpha(x_1+x_2))$$
$$= \alpha (5x_1, 0, x_1+x_2)$$
$$\therefore \boxed{L(\alpha x) = \alpha L(x)}$$

Thus the scaling property also holds!

Hence L is a Linear Transformation.

Example of Non Linear map.

The function $F: R \longrightarrow R$ is not linear
$x \longrightarrow x^2$

Property 2: $L(\alpha x) = (\alpha x)^2 = \alpha^2 x^2 \neq \alpha L(x)$

Property 2 violated!

Properties of Linear Transformations

If $L : R^m \longrightarrow R^n$ is linear, then:

- $L(0) = 0$ → 0 vector of R^n
- $L\left(\sum_{i=1}^{k} \alpha_i v_i\right) = \sum_{i=1}^{k} \alpha_i L(v_i)$, for all $\alpha_i \in R, v_i \in R^m$

0 vector of R^m

Proof:

① for linear transform: $L(v+w) = L(v) + L(w)$

$L(0+0)$ should be $L(0) + L(0)$

→ $L(0)$ should be $2L(0)$

i.e. $L(0) = 2L(0)$

→ possible only if $L(0) = 0$

②

$$L(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k) = L(\alpha_1 v_1) + L(\alpha_2 v_2) + \dots + L(\alpha_k v_k)$$
$$= \alpha_1 L(v_1) + \alpha_2 L(v_2) + \dots + \alpha_k L(v_k)$$
$$= \sum_{i=1}^{k} \alpha_i L(v_i)$$

Properties: Composition

If $L: R^m \longrightarrow R^n$ and $M: R^n \longrightarrow R^k$ are both linear, then the composite function

$$M \circ L : R^m \longrightarrow R^k$$
$$v \longrightarrow M(L(v))$$

is also Linear.

Proof:

① Let $v, w \in \mathbb{R}^m$

$$M \circ L(v+w) = M(\underline{L(v+w)})$$

$$\downarrow \text{If } L \text{ is linear}$$

$$= \underbrace{M\big(L(v)+L(w)\big)}_{\downarrow \text{ If } M \text{ is linear}}$$

$$= M\big(L(v)\big) + M\big(L(w)\big)$$

② $M \circ L(\alpha v) = M\big(L(\alpha v)\big)$

$$\downarrow \text{If } L \text{ is linear}$$

$$= M\big(\alpha\, L(v)\big)$$

$$\downarrow \text{If } M \text{ is linear}$$

$$= \alpha\, M\big(L(v)\big)$$

Matrices

Linear Maps & Matrices definition:

* Let $L : R^m \to R^n$ be a linear transformation
* Let $(e_1, \dots, e_m)$ be the canonical basis of R^m.

Then, for all $x = (x_1, \dots, x_m) \in R^m$:

$$L(x) = L\Big(\sum_{i=1}^{m} x_i e_i\Big) = \sum_{i=1}^{m} x_i\, L(e_i)$$

Thus we only need to know how to transform the basis in order to transform any vector in the vector space.

Conclusion: if you give me the vectors $L(e_1), L(e_2) \dots L(e_m) \in R^n$ then, I am able to compute $L(x)$ for any $x \in R^m$.

One needs $n \times m$ numbers to store the linear map L on a computer

This is what leads us to define a matrix!

Matrices

Defn: A $n \times m$ matrix is an array (of real numbers) with n rows and m columns. We denote by $\mathbb{R}^{n \times m}$ the set of all $n \times m$ matrices.

Canonical matrix of a Linear Map

We can encode a linear map $L: \mathbb{R}^m \to \mathbb{R}^n$ by a $n \times m$ matrix

Defn: The canonical matrix of L is the $n \times m$ matrix (that we will write also L) whose columns are $L(e_1), \dots, L(e_m)$:

$$L = \begin{pmatrix} | & | & & | \\ L(e_1) & L(e_2) & \cdots & L(e_m) \\ | & | & & | \end{pmatrix} = \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,m} \\ L_{2,1} & L_{2,2} & \cdots & L_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ L_{n,1} & L_{n,2} & \cdots & L_{n,m} \end{pmatrix}$$

where we write $L(e_j) = \begin{bmatrix} L_{1,j} \\ L_{2,j} \\ \vdots \\ L_{n,j} \end{bmatrix}$

Not we will define a few operations (so that matrices can be of use):

Matrix-vector Product

Consider a linear map $L: \mathbb{R}^m \longrightarrow \mathbb{R}^n$ and it's associated matrix $\hat{L} \in \mathbb{R}^{n \times m}$.

Question: Can we use the matrix $\hat{L}$ to compute the image $L(x)$ of a vector $x \in \mathbb{R}^m$?

Proposition:

For all $x \in \mathbb{R}^m$ we have:

$$L(x) = \hat{L}x$$

where the "matrix-vector" product $\hat{L}x \in \mathbb{R}^n$ is defined by:

$$(\hat{L}x)_i = \sum_{j=1}^{n} L_{j,i}\, x_i \quad \forall\ i \in \{1 \dots n\}$$

0

Visualizing the formula:

$L: \mathbb{R}^m \to \mathbb{R}^n$

$$(\hat{L}x)_i = \sum_{j=1}^{m} \hat{L}_{i,j} x_j = \hat{L}_{i,1} x_1 + \hat{L}_{i,2} x_2 + \dots + \hat{L}_{i,m} x_m$$

$$\begin{pmatrix} \hat{L}_{1,1} & \cdots & \hat{L}_{1,m} \\ \vdots & & \vdots \\ \hat{L}_{n,1} & \cdots & \hat{L}_{n,m} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} (\hat{L}x)_1 \\ \\ \\ \end{pmatrix} n$$

size of vector $= n$ as output space is $\mathbb{R}^n$.

Since input space is $\mathbb{R}^m$.

$$(\hat{L}x)_i = \sum_j L_{i,j} x_j = \hat{L}_{1,1}x_1 + \hat{L}_{1,2}x_2 + \dots + \hat{L}_{1,m}x_m$$

$\boxed{\text{Why do we have } L(x) = \hat{L}x \text{ ?}}$

$$\hat{L} = \begin{pmatrix} | & | & & | \\ L(e_1) & L(e_2) & \cdots & L(e_m) \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

$$= x_1 \underbrace{L(\overset{\mathbb{R}^n}{e_1})}_{\mathbb{R}^m} + x_2 L(e_2) + \dots + x_m L(e_m)$$

$$= L(x)$$

Example: Identity matrix

The identity map $\quad Id: \mathbb{R}^n \to \mathbb{R}^n, \quad x \to x \quad$ is linear.

Exercise: what is the canonical matrix of Id ?

Thus matrix will be of dimension $n \times n$

Ans

$$\widehat{Id} = \begin{pmatrix} | & | & & | \\ Id(e_1) & Id(e_2) & \cdots & Id(e_n) \\ | & | & & | \end{pmatrix}$$

$Id(e_1) = e_1$ → as the mapping is not changing anything

$Id(e_2) = e_2$

$\vdots$

$Id(e_n) = e_n$

Gives us a diagonal matrix.

Example 2: Homothety

Let $\lambda \in \mathbb{R}$. The homothety map of ratio λ:

$$H_\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
$$x \longrightarrow \lambda x$$

as we are going from $\mathbb{R}^n$ to $\mathbb{R}^n$, Homothety matrix: $n \times n$

is linear.

$v, w \in \mathbb{R}^n$

$L(v+w) = \lambda v + \lambda w$
$= L(v) + L(w)$

$L(\alpha v) = \alpha \lambda v$
$= \alpha L(v)$

Hence L is a linear map.

Exercise: what is the canonical matrix of H_λ?

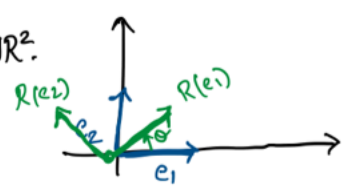
$$H_\lambda = \begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & \vdots & & \vdots \\ \vdots & 0 & \lambda & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}$$

Example 3: Rotations in $\mathbb{R}^2$.

Let $\theta \in \mathbb{R}$. The rotation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ of angle θ about origin is linear.

Exercise: What is the canonical matrix of R_θ?

Let's say we are in $\mathbb{R}^2$.



$e_1 \longrightarrow R(e_1)$

$$R(e_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \xrightarrow{R} \begin{pmatrix} 1\cos\theta \\ 1\sin\theta \end{pmatrix}$$

$$R(e_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

$$\tilde{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \end{bmatrix}$$

$$\begin{bmatrix} \sin\theta & \cos\theta \end{bmatrix}$$

Thus, $R_\theta((x,y)) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

$= \dots \Rightarrow$ co-ordinates of rotated vector.

Addition & Scalar Multiplication

$L(v+w) = L(v) + L(w)$

* Sum of two matrices of same dimension:

$$\underset{A}{\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}} + \underset{B}{\begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,m} \end{pmatrix}} = \underset{A+B}{\begin{pmatrix} a_{1,1}+b_{1,1} & \cdots & a_{1,m}+b_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1}+b_{n,1} & \cdots & a_{n,m}+b_{n,m} \end{pmatrix}}$$

* Multiplication by a scalar $\lambda \in \mathbb{R}$

$\lambda L(v) = L(\lambda v)$

$$\underset{\lambda\,(A)}{\lambda \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}} = \underset{(\lambda A)}{\begin{pmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,m} \\ \vdots & \ddots & \vdots \\ \lambda a_{n,1} & \cdots & \lambda a_{n,m} \end{pmatrix}}$$

A new vector space!

Proposition:

* $R^{n\times m}$ is a vector space
* $\dim(R^{n\times m}) = n \times m$

Proof:

$E_{i,j} \in R^{n\times m}$

$$E_{ij} = \begin{array}{c} \\ \begin{bmatrix} 0 & \cdots & & \cdots & 0 \\ \vdots & & & & \vdots \\ & & 1 & & \\ \vdots & & & & \vdots \\ 0 & \cdots & & \cdots & 0 \end{bmatrix} \\ \end{array} \begin{array}{l} \\ \\ \leftarrow i^{th} \text{ row} \\ \\ \\ \end{array}$$

j^{th} column

Verify that $E_{i,j}$ for $1 \leq i \leq n$, $1 \leq j \leq m$ is a basis

$\underline{\underline{m \times n}}$ elements

$\hookrightarrow \dim(R^{n \times m}) = n \times m$

Matrix Product

Let $M \in R^{m \times k}$ and $L \in R^{n \times m}$

$$\begin{tikzcd} \mathbb{R}^k \arrow[r, "M"] \arrow[rr, bend right, "L \circ M"'] & \mathbb{R}^m \arrow[r, "L"] & \mathbb{R}^n \end{tikzcd}$$

Definition – Proposition

* The matrix product LM is the $n \times k$ matrix of the linear map $L \circ M$

* Its coefficient are given by the formula:

$$(LM)_{i,j} = \sum_{l=1}^{m} L_{i,l} M_{l,j} \quad \text{for all} \quad 1 \le i \le n, \quad 1 \le j \le k.$$

Example: Rotations in $\mathbb{R}^2$.

The R_a and R_b denote respectively the matrices of the rotations of angles a and b about the origin, in $\mathbb{R}^2$.

Exercise: Compute the product $R_a . R_b$.

$$\underset{2\times 2}{R_a} = \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix} \qquad \underset{2\times 2}{R_b} = \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}$$

$$\underset{2\times 2}{R_a R_b} = \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix} \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}$$

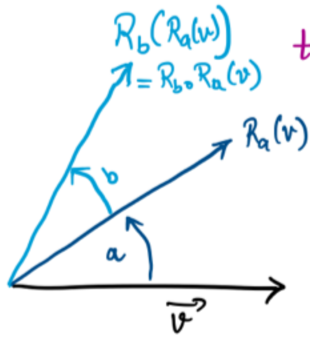
$$= \begin{bmatrix} \cos a \cos b - \sin a \sin b & -\cos a \sin b - \sin a \cos b \\ \sin a \cos b + \cos a \sin b & -\sin a \sin b + \cos a \cos b \end{bmatrix}$$

$$= \begin{bmatrix} \cos(a+b) & -\sin(a+b) \\ \sin(a+b) & \cos(a+b) \end{bmatrix}$$

Rotate by a

then Rotate by b

which confirms that transforming through R_a and then R_b would be same as transforming once through R_{a+b} !!
↳ Rotate by $a+b$



Can we divide two matrices?

For instance, if we have $AB = AC$, do we have $B = C$?

Can we do $\cancel{A}B = \cancel{A}C$?

$\boxed{\text{NO in GENERAL !}}$

↳ Exception : If A is invertible.

Invertible Matrices

A square matrix $M \in \mathbb{R}^{n\times n}$ is called invertible if there exists a matrix $M^{-1} \in \mathbb{R}^{n\times n}$ such that

$$MM^{-1} = M^{-1}M = Id_n$$

such matrix M^{-1} is unique & is called inverse of M.

Exercise : Let $A, B \in \mathbb{R}^{n\times n}$.

Show that if $AB = Id_n$ then $BA = Id_n$

Ans

$$\left.\begin{aligned} AB &= Id_n \\ ABA &= A \end{aligned}\right\} \text{Check}$$

$$ABA - A = 0$$
$$A(BA - Id) = 0$$
$$\therefore BA = Id$$

Kernel and Image

Let $L : R^m \to R^n$ be a linear transformation.

Definition (Kernel)

The kernel $Ker(L)$ (or nullspace) of L is defined as the set of all vectors $v \in \mathbb{R}^m$ such that $L(v) = 0$, i.e.

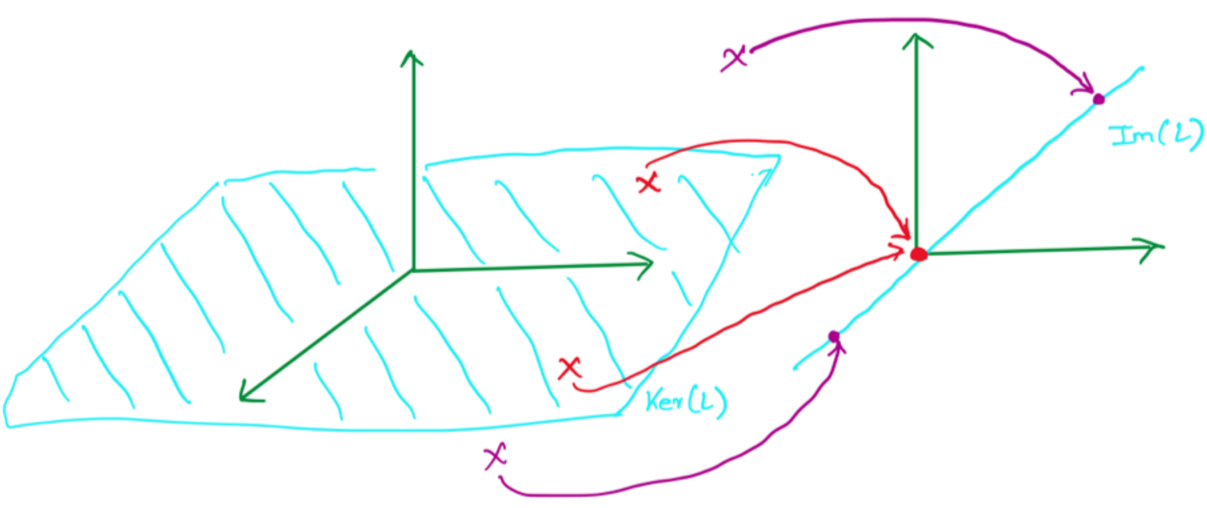
$$Ker(L) \stackrel{def}{=} \{ v \in R^m \mid L(v) = 0 \}$$

Definition (Image)

The image $Im(L)$ (or column space) of L is defined as the set of all vectors $u \in \mathbb{R}^n$ such that there exists $v \in \mathbb{R}^m$ such that $L(v) = u$.

Visualizing

Example $\mathbb{R}^3 \longrightarrow \mathbb{R}^2$



x — All points in the kernel get mapped to zero in the output space / column space / Image

x – All points outside the Kernel get mapped to some point in the Image space.
(non zero point)

Remarks

Let $L : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ be a linear transformation.

Proposition

* $\ast$ $Ker(L)$ is a subspace of $\mathbb{R}^m$.
* $\ast$ $Im(L)$ is a subspace of $\mathbb{R}^n$.

Remark : $Im(L)$ is also the span of columns of the matrix representation of L.

proof:

1 $\ast$ $Ker(L)$ is subspace of $\mathbb{R}^m$

If $v, w \in \mathbb{R}^m$ such that $\left.\begin{array}{l} L(v) = 0 \\ \& \ L(u) = 0 \end{array}\right\} \rightarrow$ i.e. $u, v \in Ker(L)$

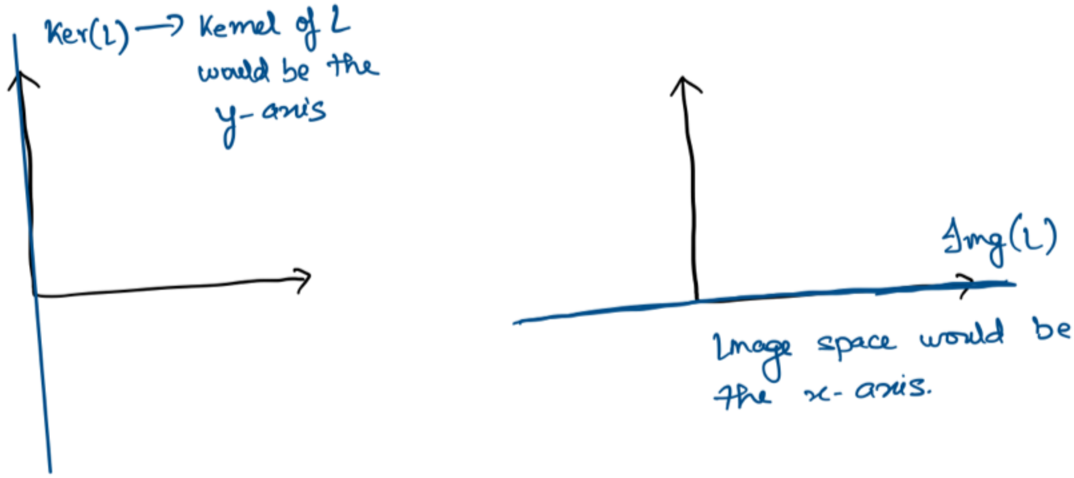
then,
$$L(v + w)$$

Example: Orthogonal Projection

Consider $L : R^2 \rightarrow R$ to be the orthogonal projection onto the x-axis.

L

$L(v) = (x, 0)$



Applications:

Solving Linear Systems

Assume that we are given a dataset:

$$a_i = (a_{i,1}, \dots, a_{i,m}) \in R^m, \quad y_i \in R \quad \text{for } i = 1, \dots, n$$

we would like to find $x \in R^m$ such that

$$x_1 a_{i,1} + \dots + x_m a_{i,m} = y_i \quad \text{for all } i = 1, \dots, n$$

This can be represented as a linear system:

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{bmatrix} \in \mathbb{R}^{n \times m} \quad \& \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

or basically,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = y_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m = y_2 \\ \\ \end{cases} \iff \underbrace{\boxed{Ax = y}}_{\text{matrix vector product}}$$

In terms of solutions to $\boxed{Ax = y}$, there will be 2 cases:
(solution is value of x)

Case 1 * $y \notin \operatorname{Im}(A)$

there cannot exist such a x

No solution.

by defn of $\operatorname{Img}(A)$, $Ax \in \operatorname{Im}(A)$
But if $y \notin \operatorname{Im}(A)$
then
no x exists to
satisfy

Case 2 : $y \in \operatorname{Im}(A)$

There will exist atleast one solution.

such that $Ax_0 = y_0$

where $x_0 \in R^m$

Rewriting:
If $y \in \operatorname{Im}(A)$ there exists x_0 such that $Ax_0 = y$

☆ Now, is this the only solution?

Let's assume x is solution,

$$\iff \quad Ax = y$$

$$\iff \quad A(x - x_0) = y - y$$

$$\iff \quad A(x - x_0) = 0$$

$$\iff \quad x - x_0 \in \operatorname{Ker}(A)$$

$\iff$ there exists $v \in Ker(A)$ such that $x = x_0 + v$

The set of all solution :

$$S = \{ x_0 + v \mid v \in Ker(A) \}$$

Conclusion :

1. $y \notin Im(A)$: there is no solution to $Ax = y$

2. $y \in Im(A)$, then there exists $x_0 \in R^m$ such that $Ax_0 = y$. The set of solutions is then

$$S = \{ x_0 + v \mid v \in Ker(A) \}$$

* If $Ker(A) = \{0\}$, then $S = \{x_0\}$: x_0 is the unique solution

* If $Ker(A) \neq \{0\}$, then $Ker(A)$ contains infinitely many vectors : there are infinitely many vectors.