★ Rank of a matrix

⇒ we define the rank of $A \in \mathbb{R}^{m \times n}$ as

$$\text{rank}(A) = \dim(\mathcal{R}(A))$$

↓

{dimention of range of A}

or

{dimention of space spand by Columns of A}

(non trivial facts)

⇒ $\text{rank}(A) = \text{rank}(A^T)$

⇒ $\text{rank}(A)$ is maximum number of independent Columns (or rows) of A.

So $\text{rank}(A) \leq \min(n, m)$

⇒ $\text{rank}(A) + \dim \mathcal{N}(A) = n$

★ Conservation of dimension

↳ or degree of freedom

$$\boxed{\text{rank}(A) + \dim \mathcal{N}(A) = n}$$

- $\text{rank}(A)$ is dimension of set 'hit' by the mapping $y = Ax$.
- $\dim \mathcal{N}(A)$ is dimension of set of x crushed to zero by $y = Ax$.

Conservation of dimension

{ Each dimension of input is either crushed to zero or ends up in output }

⇒ **Roughly Speaking**

→ n is number of degrees of freedom in input x.

→ $\dim \mathcal{N}(A)$ is number of degree of freedom lost in the mapping from x to $y = Ax$.

→ $\text{rank}(A)$ is number of degree of freedom in output y.

★ **Coding interpretation of rank**

rank of product: $\text{rank}(BC) \leq \min(\text{rank}(B), \text{rank}(C))$

⇒ If $A = BC$ with $B \in \mathbb{R}^{m \times r}$, $C \in \mathbb{R}^{r \times n}$, then $\text{rank}(A) \leq r$

⇒ Conversely, if $\text{rank}(A) = r$ then $A \in \mathbb{R}^{m \times n}$ can be factored as $A = BC$ with $B \in \mathbb{R}^{m \times r}$, $C \in \mathbb{R}^{r \times n}$.

$$x \xrightarrow{n} \boxed{C} \xrightarrow{r} \boxed{B} \xrightarrow{m} y$$

$$x \xrightarrow{n} \boxed{A} \xrightarrow{m} y$$

★ Application: fast matrix-vector multiplication

⇒ need to compute

$$y = Ax, \quad A \in \mathbb{R}^{m \times n} \qquad \text{——— (1)}$$

⇒ A has known factorization

$$A = BC, \quad B \in \mathbb{R}^{m \times r}, \quad C \in \mathbb{R}^{r \times n}$$

$$y = B(Cx) \qquad \text{——— (2)}$$

⇒ Computing y using following takes the following no of operations:

(1) mn

(2) $rn + rm = r(m+n)$

⇒ Savings can be considerable if $r \ll \min\{m, n\}$

★ Full rank matrices

⇒ We say A is full rank if $\text{rank}(A) = \min(m, n)$

- For square matrices, full rank means non singular.
- For skinny matrices $(m \geq n)$, full rank means columns are independent
- For fat matrices $(m \leq n)$, full rank means rows are independent.

[Square] [Skinny] [Fat]

★ Change of Coordinates

⇒ Standard basis vectors in R^n: $(e_1, e_2, \dots e_n)$

where,

$$e_i = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \leftarrow i^{th} \text{ row}$$

So obviously we have

$$x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

where x_i are called the coordinates of x (in the standard basis)

⇒ If $(t_1, t_2, \dots t_n)$ is another basis for R^n we have

$$x = \tilde{x}_1 t_1 + \tilde{x}_2 t_2 + \dots + \tilde{x}_n t_n$$

where $\tilde{x}_i$ are coordinates of x in the basis $(t_1, t_2 \dots t_n)$

⇒ define $T = [t_1, t_2 \dots t_n]$ so, $x = T\tilde{x}$

$$\tilde{x} = T^{-1} x$$

↳ T is invertible since t_i are a basis.

$\Rightarrow T^{-1}$ transforms (standard basis) coordinates of x into t - coordinates.

$\Rightarrow$ Inner product of i^{th} row of T^{-1} with x extracts t_i - coordinate of x.

$\Rightarrow$ Consider linear transformation $y = Ax$ $A \in R^{n \times n}$

Express y and x in terms of $t_1, t_2 \dots t_n$:

$$x = T\tilde{x} \quad , \quad y = T\tilde{y}$$

$$T\tilde{y} = AT\tilde{x}$$

$$\tilde{y} = (T^{-1}AT)\tilde{x}$$

$\Rightarrow$ $A \to T^{-1}AT$ is called Similarity transformation.

$\Rightarrow$ Similarity transformation by T expresses linear transformation $y = Ax$ in coordinates $t_1 t_2 \dots t_n$

★ (Euclidean) norm

$\Rightarrow \|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{x^T x}$

$\Rightarrow \|x\|$ measures length of vector (from origin)

$\Rightarrow$ Important properties

① $\|\alpha x\| = |\alpha| \, \|x\|$

② $\|x + y\| \leq \|x\| + \|y\|$

③ $\|x\| \geq 0$

④ $\|x\| = 0 \Leftrightarrow x = 0$

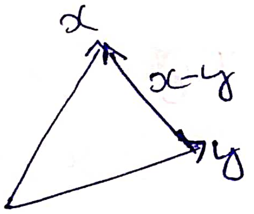
★ RMS value and (Euclidean) distance

⇒ root-mean-square (RMS) value of vector $x \in R^n$

$$\text{rms}(x) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{1/2} = \frac{\|x\|}{\sqrt{n}}$$

⇒ norm defines distance between vectors:

$$\text{dist}(x, y) = \|x - y\|$$



★ Inner Product

$$\langle x, y\rangle := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = x^T y$$

⇒ Important properties:

- $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$
- $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
- $\langle x, y\rangle = \langle y, x\rangle$
- $\langle x, x\rangle \geq 0$
- $\langle x, x\rangle = 0 \iff x = 0$

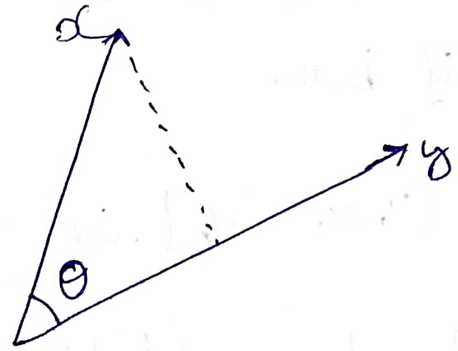
⇒ $f(y) = \langle x, y\rangle$ is linear function: $R^n \to R$, with linear map defined by row vector x^T.

★ Cauchy-Schwartz inequality and angle between vectors

⇒ for any $x, y \in R^n$, $|x^T y| \leq \|x\| \|y\|$

⇒ (Unsigned) angle between vectors in R^n defined as

$$\theta = \angle(x, y) = \cos^{-1} \frac{x^T y}{\|x\| \|y\|}$$



thus, $\boxed{x^T y = \|x\| \|y\| \cos\theta}$

- $x^T y > 0$ means $\angle(x, y)$ is acute
- $x^T y < 0$ means $\angle(x, y)$ is obtuse.

⇒ $\{x \mid x^T y \leq 0\}$ defines a half space with outward normal vector y, and boundary passing through 0.