1a) A function $T: X \to Y$ is linear if it satisfies both:

· additivity: $T(x_1+x_2) = T(x_1) + T(x_2)$

$$f(x_1+x_2) = A(x_1+x_2) = \begin{bmatrix} a_{11}(x_1+x_2)_1 + \cdots + a_{1n}(x_1+x_2)_n \\ \vdots \\ a_{m1}(x_1+x_2)_1 + \cdots + a_{mn}(x_1+x_2)_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1^1 + a_{11}x_2^1 + \cdots + a_{1n}x_1^n + a_{1n}x_2^n \\ \vdots \\ a_{m1}x_1^1 + a_{m1}x_2^1 + \cdots + a_{mn}x_1^n + a_{mn}x_2^n \end{bmatrix}$$

$$= \begin{bmatrix} (a_{11}x_1^1 + a_{12}x_1^2 + \cdots + a_{1n}x_1^n) + (a_{11}x_2^1 + a_{12}x_2^2 + \cdots + a_{1n}x_2^n) \\ \vdots \\ (a_{m1}x_1^1 + a_{m2}x_1^2 + \cdots + a_{mn}x_1^n) + (a_{m1}x_2^1 + a_{m2}x_2^2 + \cdots + a_{mn}x_2^n) \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1^1 + a_{12}x_1^2 + \cdots + a_{1n}x_1^n \\ \vdots \\ a_{m1}x_1^1 + a_{m2}x_1^2 + \cdots + a_{mn}x_1^n \end{bmatrix} + \begin{bmatrix} a_{11}x_2^1 + a_{12}x_2^2 + \cdots + a_{1n}x_2^n \\ a_{m1}x_2^1 + a_{m2}x_2^2 + \cdots + a_{mn}x_2^n \end{bmatrix}$$

$= Ax_1 + Ax_2 = f(x_1) + f(x_2)$ $\therefore$ additivity holds

· Scalar multiplication: $T(\alpha x) = \alpha T(x)$

$$f(\alpha x) = \begin{bmatrix} \alpha a_{11}x_1 + \cdots + \alpha a_{1n}x_n \\ \vdots \\ \alpha a_{m1}x_1 + \cdots + \alpha a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \alpha(a_{11}x_1 + \cdots + a_{1n}x_n) \\ \vdots \\ \alpha(a_{m1}x_1 + \cdots + a_{mn}x_n) \end{bmatrix} = \alpha \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}$$

$= \alpha Ax = \alpha f(x)$ $\therefore$ Scalar multiplication holds

Since additivity and scalar multiplication hold, it is proven that $f: \mathbb{R}^n \to \mathbb{R}^m$ is a linear funcion.

1b) $e_i \in$ Standard basis

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

$$y = f(x) = f(x_1 e_1) + f(x_2 e_2) + \cdots + f(x_n e_n) = x_1 f(e_1) + x_2 f(e_2) + \cdots + x_n f(e_n)$$

$$= x_1 \begin{bmatrix} y_{11} \\ y_{21} \\ y_{m1} \end{bmatrix} + x_2 \begin{bmatrix} y_{12} \\ y_{22} \\ \vdots \\ y_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} y_{1n} \\ y_{2n} \\ \vdots \\ y_{mn} \end{bmatrix} = \begin{bmatrix} x_1 y_{11} + x_2 y_{12} + \cdots + x_n y_{1n} \\ x_1 y_{21} + x_2 y_{22} + \cdots + x_n y_{2n} \\ x_1 y_{m1} + x_2 y_{m2} + \cdots + x_n y_{mn} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = Ax$$

c) $Ax = Bx$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \to \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ \vdots \\ b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mn}x_n \end{bmatrix}$$

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mn}x_n \end{bmatrix} = 0 \to \begin{bmatrix} (a_{11}-b_{11})x_1 + (a_{12}-b_{12})x_2 + \cdots + (a_{1n}-b_{1n})x_n \\ \vdots \\ (a_{m1}-b_{m1})x_1 + (a_{m2}-b_{m2})x_2 + \cdots + (a_{mn}-b_{mn})x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (i)$$

$$\to \begin{bmatrix} (a_{11}-b_{11}) & (a_{12}-b_{12}) & \cdots & (a_{1n}-b_{1n}) \\ \vdots & & \ddots & \vdots \\ (a_{m1}-b_{m1}) & (a_{m2}-b_{m2}) & \cdots & (a_{mn}-b_{mn}) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0 \to Cx = 0$$

For $Cx = 0 \ \forall x \in \mathbb{R}^n$, it is required that C is the zero matrix as x is not all zeros. This can be seen in (i) as we have $\sum_{i=0} (a_{ki} - b_{ki}) x_i = 0$ for $k = 1, 2, \ldots, m$, which only holds for all $(a_{ki} - b_{ki}) = 0 \to a_{ki} = b_{ki}$, which says all entries in A are the same entries in B, which is the condition for $A = B$.

2a) $\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_k a_k = 0 \to \alpha_i = 0$ for $i = 1, 2, \ldots, k$

$$\beta_1 c_1 + \beta_2 c_2 + \cdots + \beta_k c = 0 \to \beta_1 \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + \beta_2 \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} + \cdots + \beta_k \begin{bmatrix} a_k \\ b_k \end{bmatrix} = 0$$

$$\to \begin{bmatrix} \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_k a_k \\ \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_k b_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

it is already given that $\beta_i = 0$ for $i = 1, 2, \ldots, k$ in the first row, and since no component of b can contribute to the first entry, it is proven that $c_1 \cdots c_k$ are independent when $a_1 \cdots a_k$ are independent. And that the dependence of $b_1 - b_k$ is irrelevant.

b) Using the results of 2a), we cannot make this conclusion. 2a) tells us that if one of the set of vectors in $c_1 - c_k$ are independent, then so is $c_1 - c_k$. So, in this case, if $b_1 - b_k$ is independent, then so will $c_1 - c_k$ be. (by the same argument in 2a)).

$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_k a_k = 0 \rightarrow \alpha_i = 0$ and $\gamma_1 b_1 + \gamma_2 b_2 + \cdots + \gamma_k b_k = 0$ for not all $\gamma_i = 0$ $\forall i = 1,2,\ldots,k$

from (a)

$$\begin{bmatrix} \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_k a_k \\ \beta_1 b_2 + \beta_2 b_2 + \cdots + \beta_k a_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$a_1, a_2, \ldots, a_k$ are dependent so $\beta_i \neq 0$ $\forall i = 1,2,\ldots,k$

if $b_1, b_2, \ldots, b_k$ are indepent, we require $\beta_i = 0 \forall i = 1,2,\ldots,k$

Which contradicts the dependence of $a_1 - a_k$.

$\therefore$ the dependence of $a_1 - a_k$ does not gaurentee the dependence of $C_1 - C_k$.

3a) ℓ_1 norm: $\|X\|_1 = \sum_{k=1}^{n} |X_k|$

1) Nonnegativity:

$$0 \leq \|X\|_1 = |X_1| + |X_2| + \cdots + |X_n|$$

the ℓ_1 norm is a sum of non-negative values, $\therefore$ it is greater than or equal to 0.

2) $\|X\|_1 = 0$ iff $X = 0$

if $X = 0$: $\|X\|_1 = |0| + |0| + \cdots + |0| = 0$ $\therefore$ if $X = 0$, $\|X\|_1 = 0$

if $\|X\|_1 = 0$: $0 = |X_1| + |X_2| + \cdots + |X_n|$ $\rightarrow$ Since all X_i are +ve values, $\|X\|_1 = 0$ requires $X_i = 0$ $\forall i = 1,2,\ldots,n$

$\therefore$ the ℓ_1 norm is 0 if and only if $X = 0$

3) Scalar multiplication: $(\|\alpha X\|_1 = |\alpha| \|X\|_1)$

$$\|\alpha X\|_1 = |\alpha X_1| + |\alpha X_2| + \cdots + |\alpha X_n| = |\alpha||X_1| + |\alpha||X_2| + \cdots + |\alpha||X_n| = |\alpha|(|X_1| + |X_2| + \cdots + |X_n|) = |\alpha|\|X\|_1$$

$\therefore \|\alpha X\|_1 = |\alpha|\|X\|_1$

4) Triangle inequality: $(\|X+Y\|_1 \leq \|X\|_1 + \|Y\|_1)$

$$\|X+Y\| = |X_1 + Y_1| + |X_2 + Y_2| + \cdots + |X_n + Y_n| = |X_1| + |Y_1| + |X_2| + |Y_2| + \cdots + |X_n| + |Y_n|$$

$$= |X_1| + |X_2| + \cdots + |X_n| + |Y_1| + |Y_2| + \cdots + |Y_n| = \|X\|_1 + \|Y\|$$

Since $\|X+Y\|_1 = \|X\|_1 + \|Y\|_1$, the ℓ_1 norm satisfies the triangle inequality

As properties 1-4 hold, the ℓ_1 norm is indeed a norm.

b) ℓ_∞ norm: $||x||_\infty = \max_{k=1,2,\dots,n} |x_k|$

1) Nonnegativity: $||x||_\infty = |x_k| \geq 0$ due to the abs. val.

2) if $x=0$: $||x||_\infty = \max\{|0|,|0|,\dots,|0|\} = 0$ $\therefore$ if $x=0$, $||x||_\infty = 0$

if $||x||_\infty = 0 \to 0 = \max_{k=1,2,\dots,n} |x_k|$, for 0 to be max of nonnegative values, all values must be 0. $\therefore$ if $||x||_\infty = 0$, $x=0$

the ℓ_∞ norm is 0 if and only if $x=0$

3) Scalar multiplication:

$||\alpha x||_\infty = \max_{k=1,2,\dots,n} |\alpha x_k| = |\alpha x_k| = |\alpha|\,|x_k| = |\alpha|\,||x_\infty||$ $\therefore ||\alpha x||_\infty = |\alpha|\,||x||_\infty$

4) triangle inequality:

$||x+y||_\infty = \max_{k=1,2,\dots,n} |x_k+y_k| = |x_k+y_k| = |x_k| + |y_k|$ $\quad ||x||_\infty = \max_{k=1,2,\dots,n} |x_k| = |x_k|$ $\quad ||y||_\infty = \max_{k=1,2,\dots,n} |y_k| = |y_k|$

$||x+y||_\infty \leq ||x||_\infty + ||y||_\infty = |x_k| + |y_k| =$

$\therefore$ ℓ_∞ norm satisfies the triangle inequality: $||x+y||_\infty \leq ||x||_\infty + ||y||_\infty$

4a) $\langle x,x\rangle \geq 0$; $\langle x,x\rangle = 0$ iff $x=0$; $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$; $\langle \alpha x,y\rangle = \alpha\langle x,y\rangle$; $\langle x,y\rangle = \overline{\langle y,x\rangle}$

$f(x,y) = \sum_{i=1}^{n} \alpha_i x_i y_i$

1. Symmetry: no restriction

$f(x,y) = \sum_{i=1}^{n} \alpha_i x_i y_i = \sum_{i=1}^{n} \alpha_i y_i x_i = f(y,x)$

2. Positive definiteness:

$f(x,x) = \sum_{i=1}^{n} \alpha_i x_i^2 \geq 0$ if $\alpha_i \geq 0$ $\forall i=1,2,\dots,n$ $\therefore$ require $\alpha_i \geq 0$ $\forall i=1,2,\dots,n$

3. $f(x,x)=0$ if and only if $x=0$

$f(0,0) = \sum_{i=1}^{n} \alpha_i 0^2 = 0$ no restriction

$f(x,x) = 0 = \sum_{i=1}^{n} \alpha_i x_i^2$ a) for any $\alpha_i > 0$, $x_i = 0$ b) for any $\alpha_i = 0$, x can be anything $\therefore$ require $\alpha_i \neq 0$ $\forall i=1,2,\dots,n$

4. Linearity in first argument:

$f(cx+y,z) = \sum_{i=1}^{n} \alpha_i (cx_i+y_i)(z_i) = \sum_{i=1}^{n} \alpha_i c x_i z_i + \alpha_i y_i z_i = \sum_{i=1}^{n} \alpha_i c x_i z_i + \sum_{i=1}^{n} \alpha_i y_i z_i = c\sum_{i=1}^{n} \alpha_i x_i z_i + \sum_{i=1}^{n} \alpha_i y_i z_i$

$= c f(x,z) + f(y,z)$ $\therefore$ no additional restrictions necessary

$\therefore$ Under the conditions that $\alpha_i > 0$ $\forall i = 1,2,\dots,n$ for $f(x,y) = \sum_{i=1}^{n} \alpha_i x_i y_i$ to define an inner product on $\mathbb{R}^n$.

5. $\|X\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = 1 = \sum_{i=1}^{n} x_i^2 \qquad \|y\|_2 = \sqrt{\sum_{i=1}^{n} y_i^2} = 1 = \sum_{i=1}^{n} y_i^2 =$

$\langle X-y, X+y\rangle = \sum_{i=1}^{n}(x_i-y_i)(x_i+y_i) = \sum_{i=1}^{n} x_i^2 - y_i^2 = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2 = \|x\|_2 - \|y\|_2 = 1-1=0$

Since the innerproduct between (X-Y) and (X+Y) is 0, (X-Y) and (X+Y) are orthogonal

$(X-Y)+(X+Y) = 2X \to X = \frac{X+Y+X-Y}{2}$
$(X+Y)-(X-Y) = 2Y \to Y = \frac{X+Y-(X-Y)}{2}$
} We have shown that X and Y can be expressed as a linear combination of (X+Y) and (X-Y), meaning $\text{Span}(X,Y) = \text{Span}(X+Y, X-Y)$

Since X+Y and X-Y have been shown to be orthogonal, and X and Y can be expressed as a linear combination of (X+Y) and (X-Y), $B = \{X+Y, X-Y\}$ forms an orthogonal basis for the subspace spanned by X and Y.

6a) $\underset{1}{\frac{1}{\sqrt{n}}\|X\|_2 \le} \|X\|_\infty \underset{2}{\le} \|X_2\| \underset{3}{\le} \|X\|_1 \underset{4}{\le} \sqrt{n}\|X\|_2 \underset{5}{\le} n\|X\|_\infty$

1) $\|X\|_\infty = \max|x_i|$

$|x_i| \le \|x\|_\infty \to \sum_{i=1}^{n} x_i^2 \le \sum_{i=1}^{n}\|x\|_\infty^2 \to \sqrt{\sum_{i=1}^{n} x_i^2} \le \sqrt{\sum_{i=1}^{n}\|x\|_\infty^2} \to \|x\|_2 \le \sqrt{n\cdot\|x\|_\infty^2} \to \|x\|_2 \le \sqrt{n}\|x_\infty\|$

$\therefore \frac{1}{\sqrt{n}}\|x\|_2 \le \|x\|_\infty$

2) $\|X\|_\infty = \max|x_i| = x_k \qquad \|x\|_\infty^2 = x_k^2$

$\|X\|_2^2 = x_1^2 + x_2^2 + \cdots + x_k^2 + \cdots + x_n^2 \ge x_k^2$
$\ge \|x\|_\infty^2 \to \|x\|_\infty^2 \le \|x\|_2^2 \to \therefore \|x\|_\infty \le \|x\|_2$

3) $\|x\|_2^2 = \sum_{i=1}^{n} x_i^2$

terms $j=i$ — terms $i\ne j$

$\|x\|_1^2 = \left(\sum_{i=1}^{n}|x_i|\right)^2 = \sum_{j=1}^{n}\sum_{i=1}^{n}|x_i||x_j| = \sum_{i=1}^{n}\left(|x_i|^2 + 2\sum_{1\le i<j\le n}|x_i||x_j|\right) \Rightarrow \sum_{i=1}^{n} x_i^2 + 2\sum_{1\le i<j\le n}|x_i||x_j|$

$\|x\|_1^2 = \|x\|_2^2 + \{\text{Positive term since abs. val of } x_i, x_j \text{ used}\} \to \|x_1\|^2 \ge \|x_2\|^2$

$\therefore \|x_1\| \ge \|x_2\|$

4) $|x^T y| \leq \|x\|_2 \|y_2\|$ let $y = \begin{bmatrix} Sign(x_1) \\ Sign(x_2) \\ \vdots \\ Sign(x_n) \end{bmatrix}$ where $Sign(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ -1 \text{ if } x < 0 \end{cases}$

$|x^T y| = \sum_{i=1}^{n} x_i \cdot Sign(x_i) = \sum_{i=1}^{n} |x_i| = \|x\|_1$ $\|y\|_2 = \sqrt{\sum_{i=1} Sign^2(x_i)}$ Since $Sign(x_i) \in \{-1, 1\}$

$\|y\|_2 = \sqrt{n}$ $Sign(x_i)^2 = 1 \ \forall \ i = 1, 2, \dots, n$

$\therefore |x^T y| \leq \|x\|_2 \|y\|_2 \rightarrow \|x\|_1 \leq \sqrt{n} \|x\|_2$

5) from 6a.1) we know $\frac{1}{\sqrt{n}} \|x\|_2 \leq \|x\|_p \rightarrow \frac{n}{\sqrt{n}} \|x\|_2 \leq n \|x\|_p \rightarrow \sqrt{n} \|x\|_2 \leq n \|x\|_p$

$\therefore \sqrt{n} \|x\|_2 \leq n \|x\|_p$

$\therefore$ The inequality has been proven true.

6b) from 6a.5) we have $\|x\|_1 \leq \sqrt{n} \|x\|_2 \rightarrow \frac{\|x\|_1}{\|x\|_2} \leq \sqrt{n} \rightarrow n \geq \frac{\|x\|_1^2}{\|x\|_2^2}$

where n is the number of non zero inputs in y.

$\therefore Card(x) = n \geq \frac{\|x\|_1^2}{\|x\|_2^2}$

7a) $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $z_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $z_2 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$

Distance:

$z_1: \|x - z_1\| = \sqrt{(1-1)^2 + (0-1)^2} = 1$ $z_2 = \|x - z_2\| = \sqrt{(1-10)^2 + (0-0)^2} = 9$

z_1 is distance nearest neighbour. Since $z_1 < z_2$

Angle:

$z_1: \theta_1 = \cos^{-1}\left(\frac{x^T \cdot z_1}{\|x\| \|z_1\|}\right) = \cos^{-1}\left(\frac{1 \cdot 1 + 0 \cdot 1}{\sqrt{1^2+0^2} \cdot \sqrt{1^2+1^2}}\right) = 45°$ z_2 is angle nearest neighbour

$z_2: \theta_2 = \cos^{-1}\left(\frac{1 \cdot 10 + 0 \cdot 1}{\sqrt{10^2+0^2}\sqrt{1^2+0^2}}\right) = 0°$ $\because \theta_2 < \theta_1$

b) $\|x - z_i\|^2 = \|x\|^2 - 2x^T z_i + \|z_i\|^2 = \|x\|^2 - \|x\| \cancel{\|z_i\|} \cos\theta_i + 1 = \|x\|^2 - \|x\| \cos\theta_i + 1$

Since $\|x\|^2$, $\|x\|$, and 1 are constants with respect to z_i, minimizing $\|x - z_i\|^2$ is the same as maximizing $\cos\theta_i$, which is the same as minimizing θ_i

Since $\cos^{-1}(x)$ is decreasing for $x \in [-1, 1]$: if $\cos\theta_i > \cos\theta_j$, $\theta_i < \theta_j$

$\searrow$ maximizing $\cos\theta_i$, minimizes θ_i

$\therefore$ it has been shown that the vector that minimizes $\|x - z_i\|$ also minimizes θ_i resulting in the distance nearest neighbour and angle nearest neighbour being the same when z_i normalized.