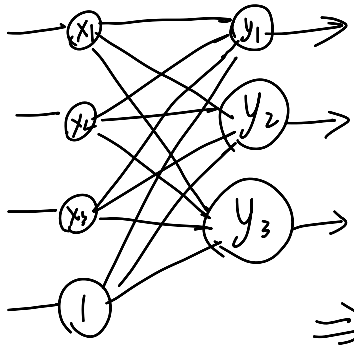
Affine space and its generation



W_{11}: weight from x_1 to y_1, b: weight from 1 to y_1

$$y_1 = W_{11}x_1 + W_{21}x_2 + W_{31}x_3 + b_1$$
$$y_2 = W_{12}x_1 + W_{22}x_2 + W_{32}x_3 + b_2$$
$$y_3 = W_{13}x_1 + W_{23}x_2 + W_{33}x_3 + b_3$$

$\Rightarrow$ compact form: $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \underbrace{\begin{bmatrix} W_{11} & W_{21} & W_{31} \\ W_{12} & W_{22} & W_{32} \\ W_{13} & W_{23} & W_{33} \end{bmatrix}}_{W} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}}_{\vec{b}}$

$\Rightarrow \vec{y} = W\vec{x} + \vec{b} \rightarrow$ affine mapping.

linear mapping

1. Group (群)

Cartansian product

Consider a set G and an operation $\otimes: G \times G \mapsto G$ defined on G.

Then, $G: (G, \otimes)$ is called a group if the following holds.

for all

① $\forall x, y \in G$, $x \otimes y \in G$ The set is closed w.r.t. the operations

② $x, y, z \in G$, $(x \otimes y) \otimes z = x \otimes (y \otimes z)$ Associativity

within

③ $\exists e \in G$, s.t. $\forall x \in G$, $x \otimes e = e \otimes x = x$ Identity element

④ $\forall x \in G$, $\exists y \in G$, s.t. $x \otimes y = y \otimes x = e$, Every element has its inverse in its group.

Specifically, G is a Abelian Group, if

① G is a group; ② $\forall x, y \in G$, $x \otimes y = y \otimes x$ commutativity

Example $(\mathbb{R} \setminus \{0\}, \cdot)$ is Abelian.

① $\forall a, b \in \mathbb{R} \setminus \{0\}$, $a \cdot b \in \mathbb{R} \setminus \{0\}$

② $\forall a, b, c \in \mathbb{R} \setminus \{0\}$, $a \cdot b \cdot c = a \cdot (b \cdot c)$

③ consider $1 \in \mathbb{R} \setminus \{0\}$, $\forall a \in \mathbb{R} \setminus \{0\}$, $a \cdot 1 = 1 \cdot a = a$

④ $\forall a \in \mathbb{R} \setminus \{0\}$, we have $a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1$

⑤ $\forall a, b \in \mathbb{R} \setminus \{0\}$, $a \cdot b = b \cdot a$

Example : $(\mathbb{R}, \cdot)$ is not a group

④ : for $0 \in \mathbb{R}$, $\nexists\, a \in \mathbb{R}$ $a \cdot 0 = 1$

Exercise: $(\mathbb{R}, +)$ is an Abelian Group

① $\forall a, b \in \mathbb{R}$, $a+b \in \mathbb{R}$

② $\forall a, b, c \in \mathbb{R}$, $a+b+c = a+(b+c)$

③ consider $0 \in \mathbb{R}$, $\forall a \in \mathbb{R}$, $a+0=0+a=a$

④ $\forall a \in \mathbb{R}$, we have $a+(-a) = -a+a = 0$

⑤ $\forall a, b \in \mathbb{R}$, $a+b = b+a$

2. Vector space

A real-valued vector space $V = (\mathcal{V}, +, \cdot)$ is a set with two operations. $V \times V$ $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$

Catasian Product

1) $+ : V \times V \longrightarrow V$ 2) $\cdot : \mathbb{R} \times V \longrightarrow V$

if: 1) $(V, +)$ is a Abelian group

2) distributivity (over $\mathbb{R}$) holds

① $\forall \lambda \in \mathbb{R}$, $\vec{x}, \vec{y} \in \mathcal{V}$, $\lambda(\vec{x}+\vec{y}) = \lambda\vec{x} + \lambda \cdot \vec{y}$

② $\forall \lambda, \psi \in \mathbb{R}$, $\vec{x} \in \mathcal{V}$, $(\lambda + \psi)\vec{x} = \lambda\vec{x} + \psi \cdot \vec{x}$

3) $\forall \lambda, \psi \in \mathbb{R}$, $\vec{x} \in \mathcal{V}$: $\lambda(\psi \cdot \vec{x}) = (\lambda\psi)\vec{x}$

4) $\forall \vec{x} \in V$: $1 \cdot \vec{x} = \vec{x}$

Space: a set with special structure among element
operation

3. Vector subspace

Let ① a vector space $V = (\mathcal{V}, +, \cdot)$; ② $\mathcal{U} \subset V$ and $\mathcal{U} \neq \emptyset$ (real subset; empty set)

Then $U := (\mathcal{U}, +, \cdot)$ is called a vector subspace of V if

① $\mathcal{U}$ is a vector space with operation $+$ and $\cdot$ restricted to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$, respectively. $\Rightarrow U \subseteq V$

3. (1) Linear Combination.

Consider a vector space $V := (\mathcal{V}, +, \cdot)$ and a finite number of vectors $\vec{x}_1, \ldots, \vec{x}_k \in \mathcal{V}$. Then, every other $\vec{v} \in \mathcal{V}$ of the form

$$\vec{v} = \lambda_1 \vec{x} + \lambda_2 \vec{x}_2 + \cdots + \lambda_k \vec{x}_k = \sum_{i=1}^{k} \lambda_i \vec{x}_i \in \mathcal{V}$$

with $\lambda_1, \ldots \lambda_k \in \mathbb{R}$ is a linear combination of vectors $\vec{x}_1, \ldots \vec{x}_k$

(2) Linear dependency

Consider a vector space V, with $k > 0$, vector $\vec{x}_1, \ldots, \vec{x}_k \in V$ and $\underline{\sum_{i=1}^{k} \lambda_i \vec{x}_i = \vec{0}}$

We say $\vec{x}_1, \ldots, \vec{x}_k$ are

① Linearly independent: only trivial solution exists, i.e. $\lambda_1 = \cdots = \lambda_k = 0$

② Linearly dependent: we can find a solution with $\lambda_i \neq 0$ for some $i \in [0, k]$ $\quad \lambda_j \neq 0 \quad \sum_{i=1}^{k/\{j\}} \lambda_i \vec{x}_i = -\lambda_j \vec{x}_j$

(3) Spanning / Generating sets

Consider a vector space $V := (\mathcal{V}, +, \cdot)$ and a set $A = \{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq V$, if $\forall \vec{v} \in \mathcal{V}$, $\exists \lambda_1, \ldots, \lambda_k \in \mathbb{R}$, s.t.

for all   there exits

$$\vec{v} = \sum_{i=1}^{k} \underline{\lambda_i} \vec{x}_i \leftarrow \lambda_i : \text{coordinate.}$$

$\Rightarrow$ A is called a spanning / generating set of V, denoted by span$(A) = V$

(4) minimal generating set (a.k.a. basis)

A is the minimal generating set of V if

① A is a generating set ; ② $\nexists A' \subset A$, s.t. A' spans V

Take away:

① basis : a minimal set of linear independent vectors that spans V

② In general, vectors in a generating set could be linear dependent

③ Dimension : the number of basis vector, denoted by $\dim(V)$

Affine space and its generation 2: linear mapping and affine mapping

$$\vec{y} = \tilde{W}\vec{x} + \vec{b}$$

Another way of viewing $\tilde{W}\vec{x}$, let $\vec{x} = [x_1, x_2, x_3]^T \in V$

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_{11}x_1 + w_{12}x_2 + w_{13}x_3 \\ w_{21}x_1 + w_{22}x_2 + w_{23}x_3 \\ x_{31}x_1 + w_{32}x_2 + w_{33}x_3 \end{bmatrix}$$

$$= x_1 \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} + x_2 \begin{bmatrix} w_{12} \\ w_{22} \\ w_{32} \end{bmatrix} + x_3 \begin{bmatrix} w_{13} \\ w_{23} \\ w_{33} \end{bmatrix}$$

$\Rightarrow$ $\tilde{w}\vec{x}$ is a vector within the span of column of $\tilde{w}$

$\Rightarrow$ $\tilde{w}\vec{x}$ is a linear transformation of $[x_1, x_2, x_3]^T$ from its own space (denoted by V) to the colume space of $\tilde{w}$ $\Rightarrow$ a mapping from $V \to \text{column}(w)$ is a linear mapping.

Def (linear mapping) given two vector space V, W, a mapping $\phi: V \mapsto W$ is a linear mapping if

$$\forall \vec{x}, \vec{y} \in V, \ \forall \lambda, \psi \in \mathbb{R}: \ \phi(\lambda\vec{x} + \psi\vec{y}) = \lambda\phi(\vec{x}) + \psi\phi(\vec{y})$$

$\Rightarrow$ linear mapping preserves the structure of vector space.

$\Rightarrow$ the order between linear mapping and operations of the vector space can be exchanged.

$\tilde{w}\vec{x}$: given $\vec{x} \in V$, $\tilde{w}\vec{x}$ maps $\vec{x}$ to space W,

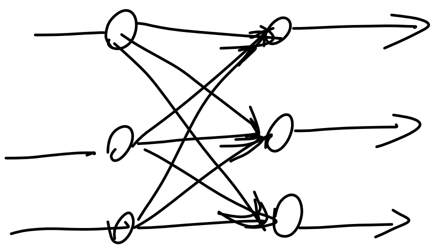
let $\vec{x} \in \mathbb{R}^n$, $\tilde{w}\vec{x} \in \mathbb{R}^m$, we denote the number of linearly independent columns of a matrix W by rank(W), or the rank of W

important property: for $\tilde{W} \in \mathbb{R}^{m \times n}$, if $\text{rank}(\tilde{W}) = \min\{m, n\}$

the we say W has full rank

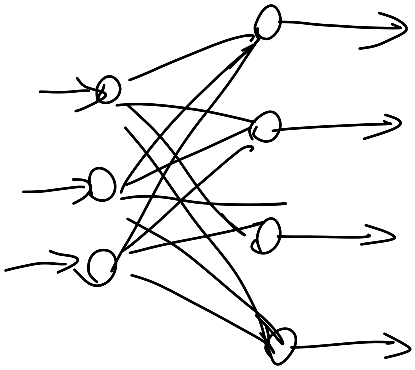
Three situation (consider $\tilde{W}$ with full rank)

① $\text{rank}(\tilde{W}) = n = m$, space V and W has the same dimension



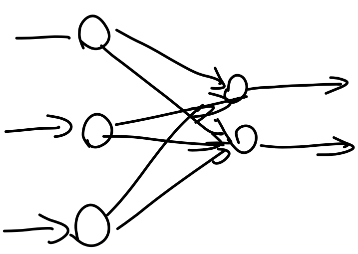
we representing a same vector using different basis.

② $\text{rank}(\tilde{W}) = n < m$, meaning that W has a higher dim. than V, we are mapping the input data to a higher dimensional space (a.k.a. upsampling)



handling data in lower dimension in a higher dimension

③ $\text{rank}(\tilde{W}) = m < n$, meaning that W has a lower dim. than V, we are mapping the input data to a lower dimensional space (a.k.a. downsampling)



dimensionality reduction

handling "the most important" pespective (lower dim.) of some high dimensional data

Exercise basis change

Consider vector space V and W, and order bases

$B = (\vec{b}_1, \dots, \vec{b}_n)$, $\tilde{B} = (\vec{b}_1, \dots, \vec{b}_n)$ of V

$C = (\vec{c}_1, \dots, \vec{c}_m)$, $\tilde{C} = (\vec{c}_1', \dots, \vec{c}_m')$ of W

Let A_ϕ, S, T being the transformation matrices from B to C, from $\tilde{B}$ to B, and from $\tilde{C}$ to C

$\Rightarrow$ what is the transformation matric from $\tilde{B}$ to $\tilde{C}$

$$\tilde{A}_\phi = ?$$

Answer

$$\begin{tikzcd} B \arrow[r, "A_\phi"] & C \arrow[d, shift right, "T^{-1}"'] \\ \tilde{B} \arrow[u, "S"] \arrow[r] & \tilde{C} \arrow[u, shift right, "T"'] \end{tikzcd}$$

consider $\vec{v} \in V$ with basis $\tilde{B}$

$\tilde{B} \to B : S\vec{v}$

$\tilde{B} \to B \to C : A_\phi S \vec{v}$

$\tilde{B} \to B \to C \to \tilde{C} : T^{-1} A_\phi S \vec{v}$

$$\Rightarrow \tilde{A}_\phi = T^{-1} A_\phi S$$

Remark 1: Matrices A_ϕ and $\tilde{A}_\phi$ are equivalent if there exists regular matrix (i.e. matrix with full rank) $S \in \mathbb{R}^{n \times n}$ and $T \in \mathbb{R}^{m \times m}$, s.t. $\tilde{A}_\phi = T^{-1} A_\phi S$

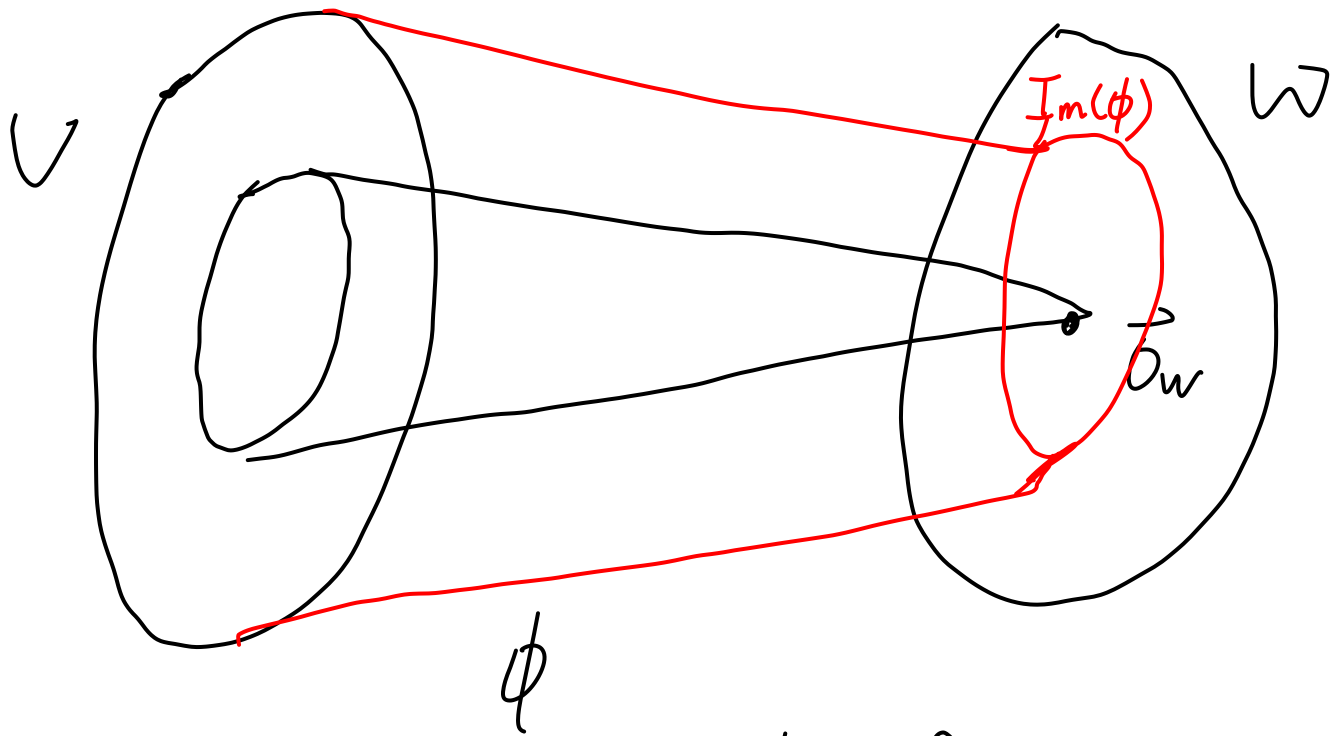
Remark 2: Matrices A_ϕ and $\tilde{A}_\phi$ are similar if there exist a regular matrix $S \in \mathbb{R}^{n \times n}$, s.t. $\tilde{A}_\phi = S^{-1} A_\phi S$

2, Image and kernel

For a mapping $\phi : V \mapsto W$ (not restricted to linear mapping)

Kernel: $\ker(\phi) := \phi^{-1}(\vec{0}_W) = \{\vec{v} \in V \mid \phi(\vec{v}) = \vec{0}_W\}$

Image: $\mathrm{Image}(\phi) := \phi(V) = \{\vec{w} \in W \mid \exists \vec{v} \in V, \text{ s.t. } \phi(\vec{v}) = \vec{w}\}$



Consider linear mapping, transformatrix matrix A

kernel: $\ker(\phi) := \{\vec{v} \in V \mid A\vec{v} = \vec{0}_W\}$ Null Space

$\Rightarrow$ capture all possible linear combination of ele. in V that produce $\vec{0}_W$

$\mathrm{Image}(\phi) = \{\vec{w} \in W \mid \exists \vec{v} \in V, \text{ s.t. } A\vec{v} = \vec{w}\}$

$\Rightarrow$ span of column vector ~~in~~ of A : column space

more properties: session 2.7.3

3. Affine spaces

Def. Let V be a vector space, $\vec{x}_0 \in V$, and $U \subseteq V$ be a subspace

We called

$$L = \vec{x}_0 + U := \{\vec{x}_0 + \vec{u} \mid \vec{u} \in U\}$$

$$= \{\vec{v} \in V \mid \exists \vec{u} \in U, \text{ s.t. } \vec{v} = \vec{x}_0 + \vec{u}\} \subseteq V$$

an affine subspace (or linear manifold) of V

$\vec{x}_0$: support point $\quad$ U: direction space

Exercise: prove: if $\vec{x}_0 \notin U$, then $\vec{0} \notin L$ ← not in

proof: Suppose $\vec{x}_0 \notin U$, but $\vec{0} \in L$, then $\exists \vec{u} \in U$, s.t. $\vec{x}_0 + \vec{u} = \vec{0}$

$\Rightarrow -\vec{x}_0 \in U$

Since U is a vector subspace, then $(U, +)$ is an Abelian group.

$\Rightarrow$ for $-\vec{x}_0 \in U$, $\underline{\exists \vec{a} \in U}$, s.t. $-\vec{x}_0 + \vec{a} = \vec{0}$

$\Rightarrow \vec{a} = \vec{x}_0 \in U$ $\quad$ ✓

4. Affine mapping

Def. Given two vector spaces V, W, a linear mapping $\phi: V \mapsto W$, $\vec{a} \in W$

the mapping $\phi: V \mapsto W$ with

$$\phi'(\vec{x}) = \vec{a} + \phi(\vec{x})$$

is called an affine mapping from V to W.

$\vec{a}$ ↑ translation vector

Summary

$\vec{y} = W\vec{x} + b$ $\quad$ Sec 2.4 ~ 2.9

Group

↳ Abelian g. → vector space

↓

space $\quad$ eneration $\begin{cases} \text{linear dependency} \\ \text{basis, generating set} \end{cases}$

↓ $W\vec{x}$

linear mapping → rank(W) $\begin{cases} \text{coordinate transformation} \\ \text{downsampling} \\ \text{upsampling} \end{cases}$

linear mapping → $\begin{cases} \text{kernel} \to \text{null space} \\ \text{image} \to \text{columnspace} \end{cases}$

↓ translation

affine mapping