3. Linear Transformation

- linear transformation
- matrix transformation
- kernel and range
- isomorphism
- composition
- inverse transformation

Transformation

let X and Y be vector spaces

a **transformation** T from X to Y, denoted by

$$T : X \to Y$$

is an assignment taking $x \in X$ to $y = T(x) \in Y$,

$$T : X \to Y, \quad y = T(x)$$

- **domain** of T, denoted $\mathcal{D}(T)$ is the collection of all $x \in X$ for which T is defined
- vector $T(x)$ is called the **image** of x under T
- collection of all $y = T(x) \in Y$ is called the **range** of T, denoted $\mathcal{R}(T)$

example 1 define $T : \mathbf{R}^3 \to \mathbf{R}^2$ as

$$\begin{aligned} y_1 &= -x_1 + 2x_2 + 4x_3 \\ y_2 &= -x_2 + 9x_3 \end{aligned}$$

where $x \in \mathbf{R}^3$ and $y \in \mathbf{R}^2$

example 2 define $T : \mathbf{R}^3 \to \mathbf{R}$ as

$$y = \sin(x_1) + x_2 x_3 - x_3^2$$

where $x \in \mathbf{R}^3$ and $y \in \mathbf{R}$

example 3 general transformation $T : \mathbf{R}^n \to \mathbf{R}^m$

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_n) \\ y_2 &= f_2(x_1, x_2, \ldots, x_n) \\ &\vdots \\ y_m &= f_m(x_1, x_2, \ldots, x_n) \end{aligned}$$

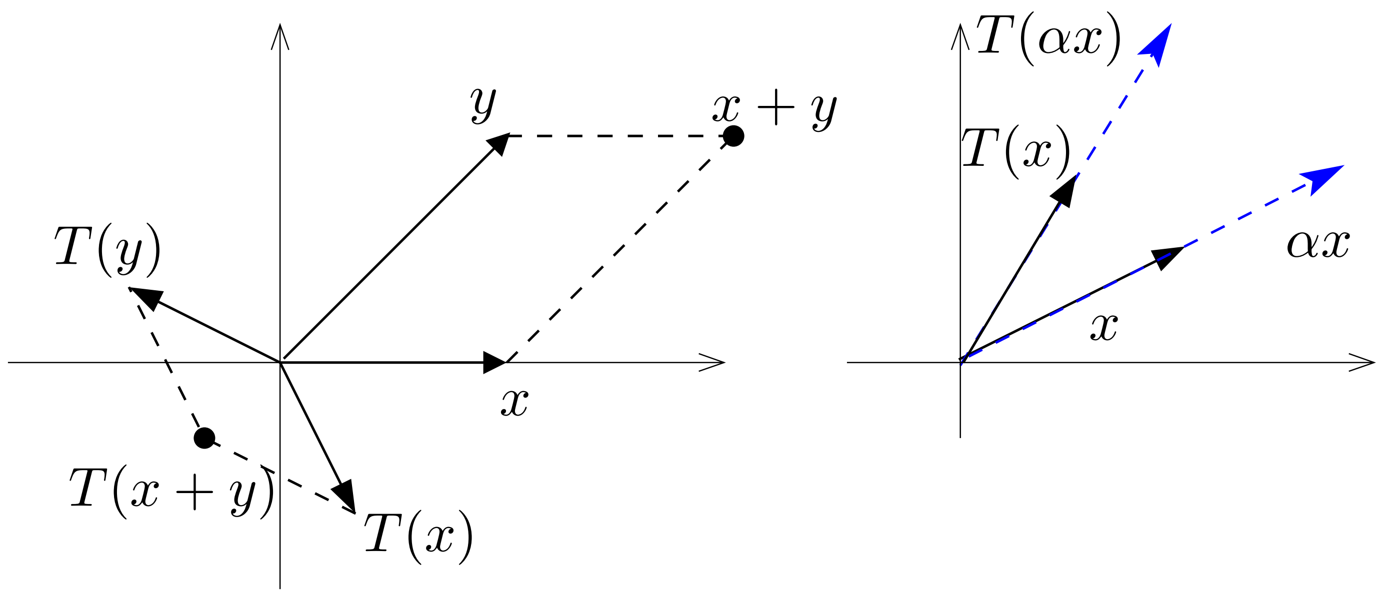
where $f_1, f_2, \ldots, f_m$ are real-valued functions of n variables

Linear transformation

let X and Y be vector spaces over **R**

Definition: a transformation $T : X \to Y$ is **linear** if

- $T(x + z) = T(x) + T(z), \quad \forall x, y \in X$ (additivity)
- $T(\alpha x) = \alpha T(x), \quad \forall x \in X, \forall \alpha \in \mathbf{R}$ (homogeneity)



Examples

✎ which of the following is a linear transformation ?

- **matrix transformation** $T : \mathbf{R}^n \to \mathbf{R}^m$

$$T(x) = Ax, \quad A \in \mathbf{R}^{m \times n}$$

- **affine transformation** $T : \mathbf{R}^n \to \mathbf{R}^m$

$$T(x) = Ax + b, \quad A \in \mathbf{R}^{m \times n}, \quad b \in \mathbf{R}^m$$

- $T : \mathbf{P}_n \to \mathbf{P}_{n+1}$

$$T(p(t)) = tp(t)$$

- $T : \mathbf{P}_n \to \mathbf{P}_n$

$$T(p(t)) = p(t+1)$$

- $T : \mathbf{R}^{m\times n} \to \mathbf{R}^{n\times m}, \quad T(X) = X^T$
- $T : \mathbf{R}^{n\times n} \to \mathbf{R}, \quad T(X) = \det(X)$
- $T : \mathbf{R}^{n\times n} \to \mathbf{R}, \quad T(X) = \mathbf{tr}(X)$
- $T : \mathbf{R}^{n} \to \mathbf{R}, \quad T(x) = \|x\| \triangleq \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$
- $T : \mathbf{R}^{n} \to \mathbf{R}^{n}, \quad T(x) = 0$

  denote $F(-\infty,\infty)$ the set of all real-valued functions on $(-\infty,\infty)$

- $T : C^1(-\infty,\infty) \to F(-\infty,\infty)$

$$T(f) = f'$$

- $T : C(-\infty,\infty) \to C^1(-\infty,\infty)$

$$T(f) = \int_0^t f(s)ds$$

Examples of matrix transformation

$T : \mathbf{R}^n \to \mathbf{R}^m$

$$T(x) = Ax, \quad A \in \mathbf{R}^{m \times n}$$

zero transformation: $T : \mathbf{R}^n \to \mathbf{R}^m$

$$T(x) = 0 \cdot x = 0$$

T maps every vector into the zero vector

identity operator: $T : \mathbf{R}^n \to \mathbf{R}^n$

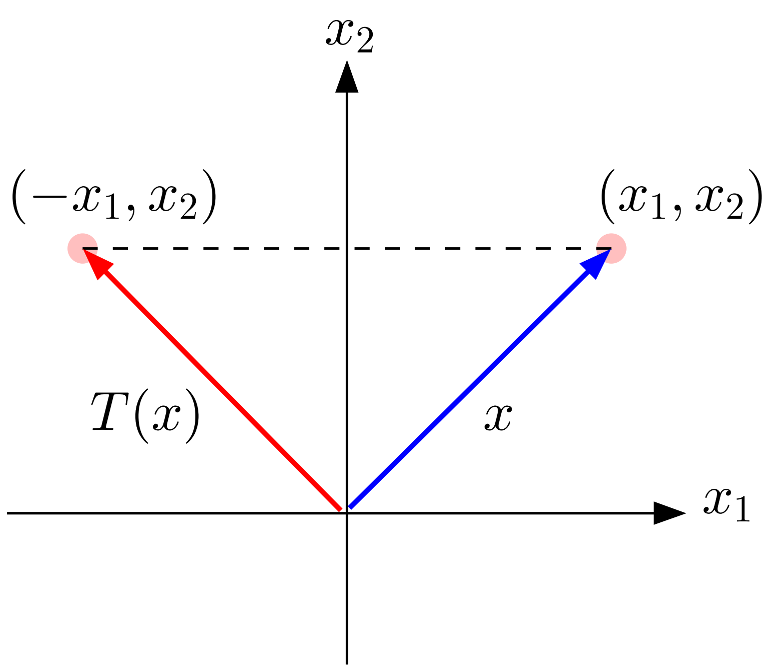
$$T(x) = I_n \cdot x = x$$

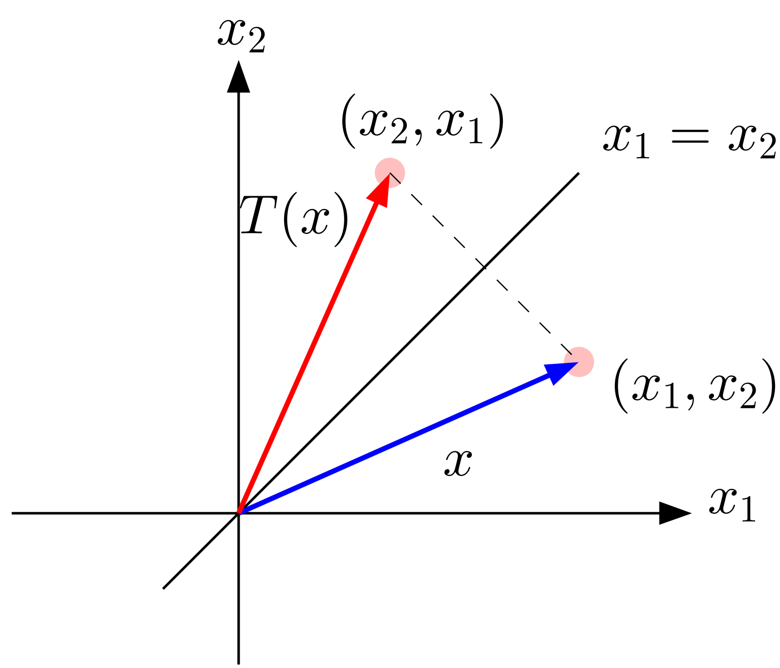
T maps a vector into itself

Geomatrix transformations

reflection operator: $T : \mathbf{R}^n \to \mathbf{R}^n$

T maps each point into its symmetric image about an axis or a line

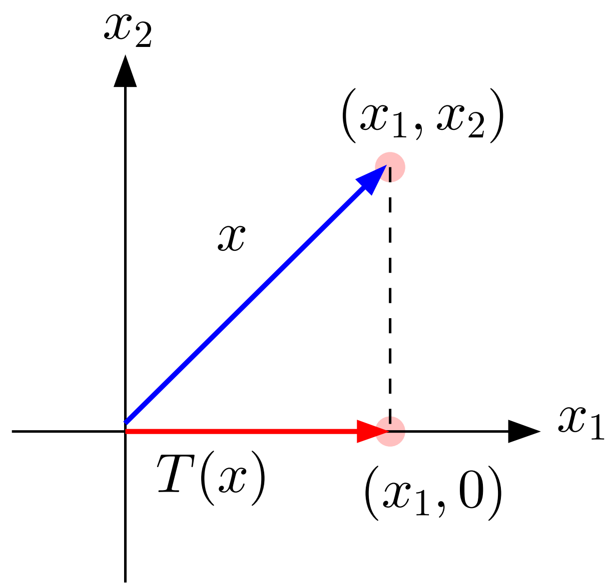


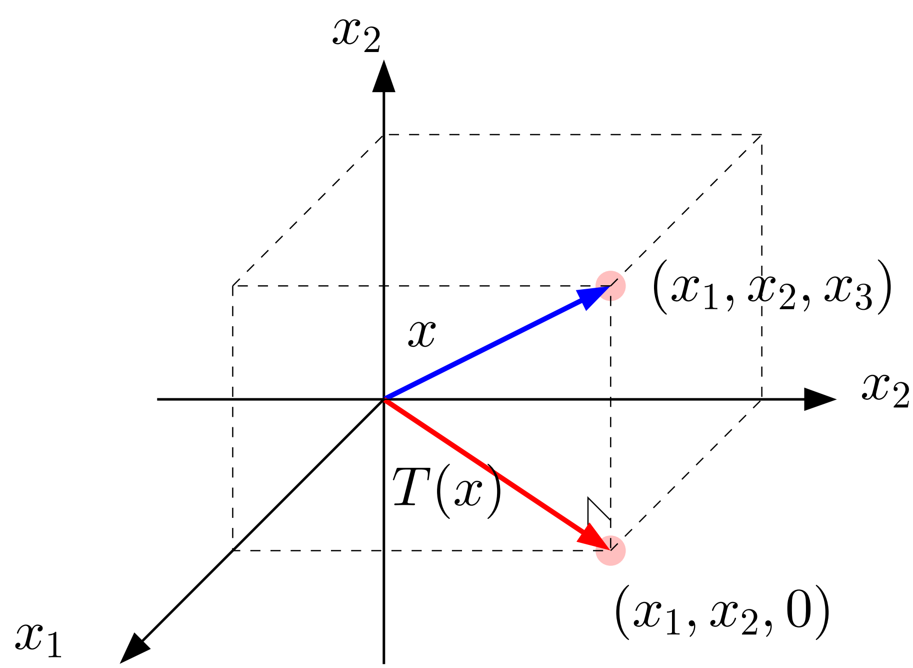


$$T(x) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} x \qquad T(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x$$

projection operator: $T : \mathbf{R}^n \to \mathbf{R}^n$

T maps each point into its orthogonal projection on a line or a plane



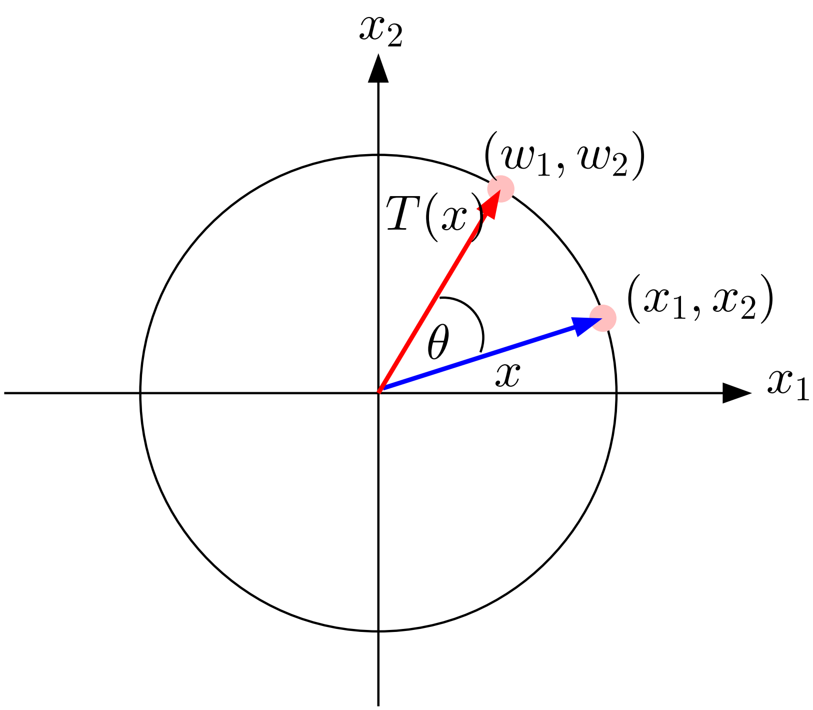


$$T(x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x$$

$$T(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} x$$

rotation operator: $T : \mathbf{R}^n \to \mathbf{R}^n$

T maps points along circular arcs

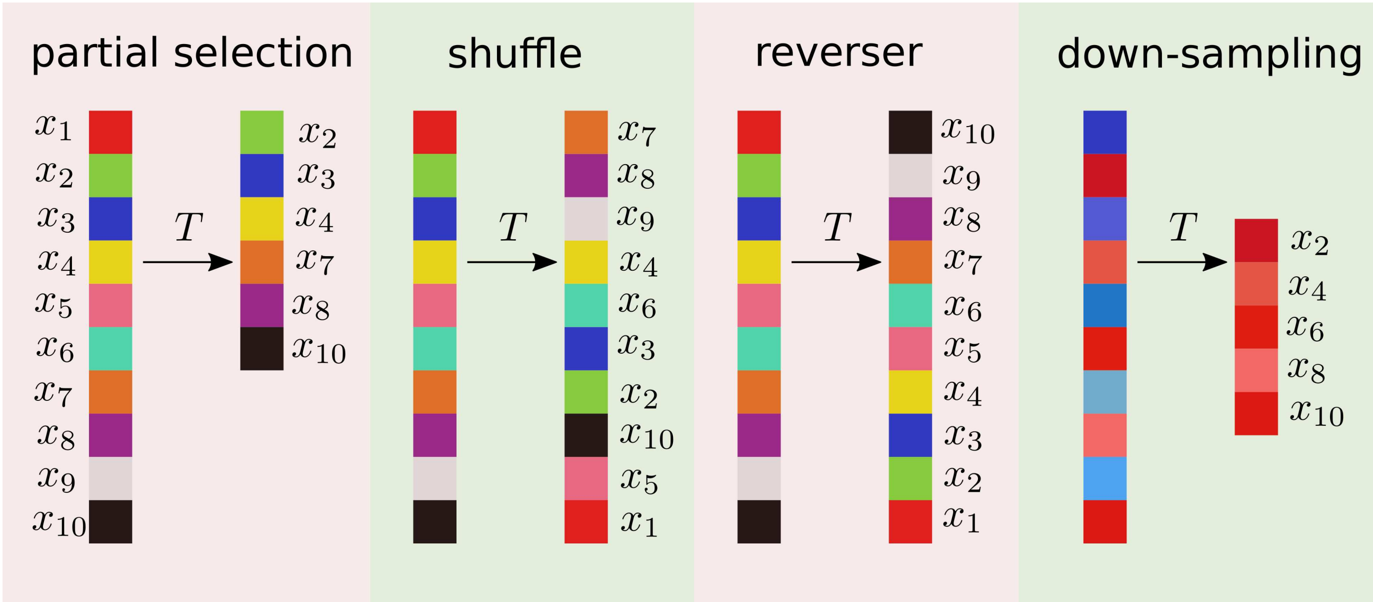


T rotates x through an angle θ

$$w = T(x) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} x$$

Selector transformations

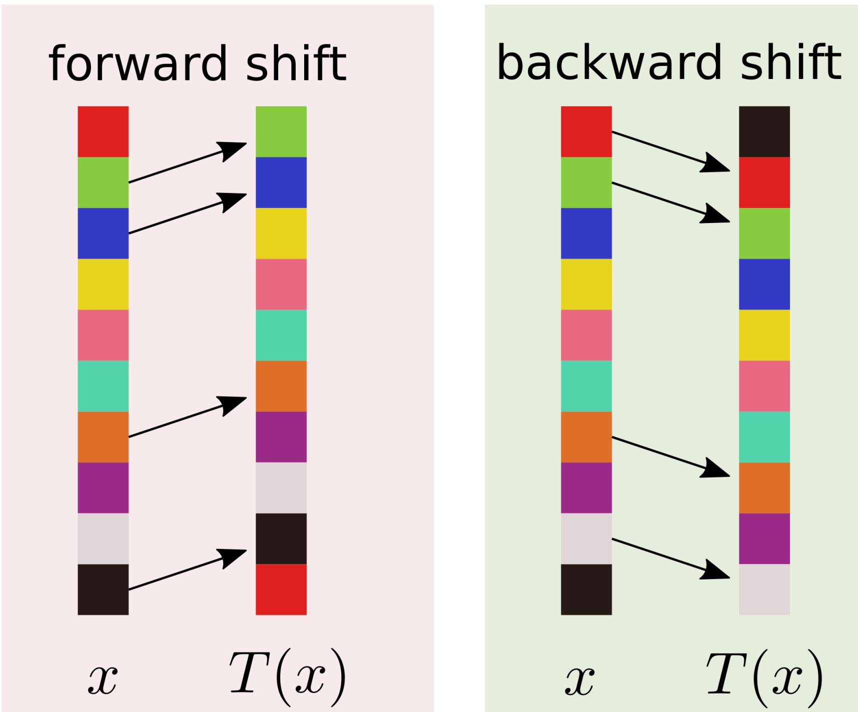
these transformations can be represented as $y = T(x) = Ax$



- partial selection: select some entries of x
- shuffle: randomize entries in x
- reverser: reverse the order of x
- down-sampling: sub-sample entries in x, *e.g.*, `x(1:2:end)`

Shift transformations

shifting sequences as a matrix transformation $T(x) = Ax$



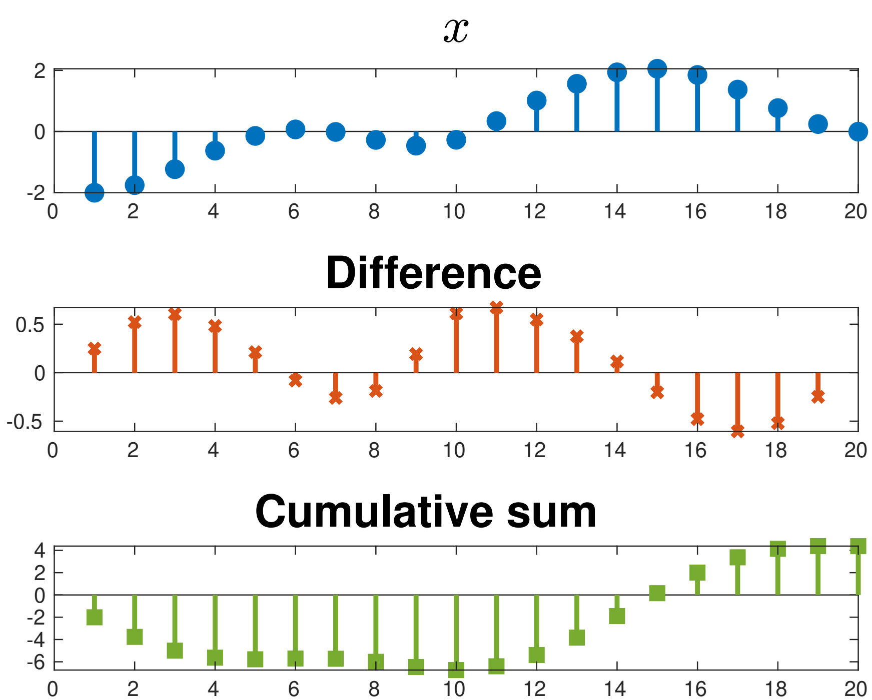
$$T_1(x) = \begin{bmatrix} x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \\ x_1 \end{bmatrix}, \quad T_2(x) = \begin{bmatrix} x_n \\ x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{bmatrix}$$

what is the associated matrix A for each transformation ?

do you notice some structure of A ?

Signal processing

differencing and cumulative sum as matrix transformations $T(x) = Ax$



$$T_1(x) = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ \vdots \\ x_n - x_{n-1} \end{bmatrix}$$

$$T_2(x) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \\ \vdots \\ x_1 + x_2 + \cdots + x_n \end{bmatrix}$$

`diff` and `cumsum` commands in MATLAB

what is the associated matrix A for each transformation ?

Image transformation

cropping a 1200×850-pixel image to 490×430-pixel image

transformation of a matrix of $M \times N$ to the size of $m \times n$

$$T : \mathbf{R}^{M \times N} \to \mathbf{R}^{m \times n}, \quad T(X) = AXB$$

where A selects the rows of X and B selects the columns of X

Image of linear transformation

let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces and a basis for $\mathcal{V}$ is

$$S = \{v_1, v_2, \ldots, v_n\}$$

let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation

the image of any vector $v \in \mathcal{V}$ under T can be expressed by

$$T(v) = a_1 T(v_1) + a_2 T(v_2) + \cdots + a_n T(v_n)$$

where $a_1, a_2, \ldots, a_n$ are coefficients used to express v, *i.e.*,

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$$

(follow from the linear property of T)

Kernel and Range

let $T : X \to Y$ be a linear transformation from X to Y

Definitions:

kernel of T is the set of vectors in X that T maps into 0

$$\mathbf{ker}(T) = \{x \in X \mid T(x) = 0\}$$

range of T is the set of all vectors in Y that are images under T

$$\mathcal{R}(T) = \{y \in Y \mid y = T(x), \quad x \in X\}$$

Theorem ✎

- $\mathbf{ker}(T)$ is a subspace of X
- $\mathcal{R}(T)$ is a subspace of Y

matrix transformation: $T : \mathbf{R}^n \to \mathbf{R}^m$, $\quad T(x) = Ax$

- $\mathbf{ker}(T) = \mathcal{N}(A)$: kernel of T is the nullspace of A
- $\mathcal{R}(T) = \mathcal{R}(A)$: range of T is the range (column) space of A

zero transformation: $T : \mathbf{R}^n \to \mathbf{R}^m$, $\quad T(x) = 0$

$$\mathbf{ker}(T) = \mathbb{R}^n, \quad \mathcal{R}(T) = \{0\}$$

identity operator: $T : \mathcal{V} \to \mathcal{V}$, $\quad T(x) = x$

$$\mathbf{ker}(T) = \{0\}, \quad \mathcal{R}(T) = \mathcal{V}$$

differentiation: $T : C^1(-\infty, \infty) \to F(-\infty, \infty)$, $\quad T(f) = f'$

$\mathbf{ker}(T)$ is the set of constant functions on $(-\infty, \infty)$

Rank and Nullity

Rank of a linear transformation $T : X \to Y$ is defined as

$$\mathbf{rank}(T) = \dim \mathcal{R}(T)$$

Nullity of a linear transformation $T : X \to Y$ is defined as

$$\mathbf{nullity}(T) = \dim \mathbf{ker}(T)$$

(provided that $\mathcal{R}(T)$ and $\mathbf{ker}(T)$ are finite-dimensional)

Rank-Nullity theorem: suppose X is a finite-dimensional vector space

$$\mathbf{rank}(T) + \mathbf{nullity}(T) = \dim(X)$$

Proof of rank-nullity theorem

- assume $\dim(X) = n$
- assume a nontrivial case: $\dim \mathbf{ker}(T) = r$ where $1 < r < n$
- let $\{v_1, v_2, \ldots, v_r\}$ be a basis for $\mathbf{ker}(T)$
- let $W = \{v_1, v_2, \ldots, v_r\} \cup \{v_{r+1}, v_{r+2}, \ldots, v_n\}$ be a basis for X
- we can show that
$$S = \{T(v_{r+1}), \ldots, T(v_n)\}$$
forms a basis for $\mathcal{R}(T)$ $\quad$ ($\therefore$ complete the proof since $\dim\ S = n - r$)

$\mathbf{span}\ S = \mathcal{R}(T)$

- for any $z \in \mathcal{R}(T)$, there exists $v \in X$ such that $z = T(v)$
- since W is a basis for X, we can represent $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$
- we have $z = \alpha_{r+1} T(v_{r+1}) + \cdots + \alpha_n T(v_n)$ $\quad$ ($\because\ v_1, \ldots, v_r \in \mathbf{ker}(T)$)

S **is linearly independent**, *i.e.*, we must show that

$$\alpha_{r+1}T(v_{r+1}) + \cdots + \alpha_n T(v_n) = 0 \quad \Longrightarrow \quad \alpha_{r+1} = \cdots = \alpha_n = 0$$

- since T is linear

$$\alpha_{r+1}T(v_{r+1}) + \cdots + \alpha_n T(v_n) = T(\alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n) = 0$$

- this implies $\alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n \in \mathbf{ker}(T)$

$$\alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n = \alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_r v_r$$

- since $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ is linear independent, we must have

$$\alpha_1 = \cdots = \alpha_r = \alpha_{r+1} = \cdots = \alpha_n = 0$$

One-to-one transformation

a linear transformation $T : X \to Y$ is said to be **one-to-one** if

$$\forall x, z \in X \qquad T(x) = T(z) \quad \Longrightarrow \quad x = z$$

- T never maps distinct vectors in X to the same vector in Y
- also known as **injective** tranformation

✌ **Theorem:** T is *one-to-one* if and only if $\mathbf{ker}(T) = \{0\}$, $i.e.$,

$$T(x) = 0 \quad \Longrightarrow \quad x = 0$$

- for $T(x) = Ax$ where $A \in \mathbf{R}^{n \times n}$,

$$T \text{ is one-to-one} \quad \Longleftrightarrow \quad A \text{ is } \textit{invertible}$$

Onto transformation

a linear transformation $T : X \to Y$ is said to be **onto** if

for **every** vector $y \in Y$, there exists a vector $x \in X$ such that

$$y = T(x)$$

- every vector in Y is the image of at least one vector in X
- also known as **surjective** transformation

✌ **Theorem:** T is onto if and only if $\mathcal{R}(T) = Y$

✌ **Theorem:** for a *linear operator* $T : X \to X$,

$$T \text{ is one-to-one if and only if } T \text{ is onto}$$

✎ which of the following is a one-to-one transformation ?

- $T : \mathbf{P}_n \to \mathbf{R}^{n+1}$

$$T(p(t)) = T(a_0 + a_1 t + \cdots + a_n t^n) = (a_0, a_1, \ldots, a_n)$$

- $T : \mathbf{P}_n \to \mathbf{P}_{n+1}$

$$T(p(t)) = tp(t)$$

- $T : \mathbf{R}^{m\times n} \to \mathbf{R}^{n\times m}, \quad T(X) = X^T$

- $T : \mathbf{R}^{n\times n} \to \mathbf{R}, \quad T(X) = \mathbf{tr}(X)$

- $T : C^1(-\infty, \infty) \to F(-\infty, \infty), \quad T(f) = f'$

Matrix transformation

consider a linear transformation $T : \mathbf{R}^n \to \mathbf{R}^m$,

$$T(x) = Ax, \qquad A \in \mathbf{R}^{m \times n}$$

✌ **Theorem:** the following statements are equivalent

- T is **one-to-one**
- the homonegenous equation $Ax = 0$ has only the trivial solution $(x = 0)$
- $\mathbf{rank}(A) = n$

✌ **Theorem:** the following statements are equivalent

- T is **onto**
- for every $b \in \mathbf{R}^m$, the linear system $Ax = b$ always has a solution
- $\mathbf{rank}(A) = m$

Isomorphism

a linear transformation $T : X \to Y$ is said to be an **isomorphism** if

$$T \text{ is both one-to-one and onto}$$

if there exists an isomorphism between X and Y, the two vector spaces are said to be **isomorphic**

✌ **Theorem:**

- for any n-dimensional vector space X, there always exists a linear transformation $T : X \to \mathbf{R}^n$ that is one-to-one and onto (for example, a coordinate map)
- every real n-dimensional vector space is isomorphic to $\mathbf{R}^n$

examples of isomorphism

- $T : \mathbf{P}_n \to \mathbf{R}^{n+1}$

$$T(p(t)) = T(a_0 + a_1 t + \cdots + a_n t^n) = (a_0, a_1, \ldots, a_n)$$

  $\mathbf{P}_n$ is isomorphic to $\mathbf{R}^{n+1}$

- $T : \mathbf{R}^{2\times 2} \to \mathbf{R}^4$

$$T\left(\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}\right) = (a_1, a_2, a_3, a_4)$$

  $\mathbf{R}^{2\times 2}$ is isomorphic to $\mathbf{R}^4$

in these examples, we observe that

- T maps a vector into its coordinate vector relative to a standard basis
- for any two finite-dimensional vector spaces that are isomorphic, they have the same dimension

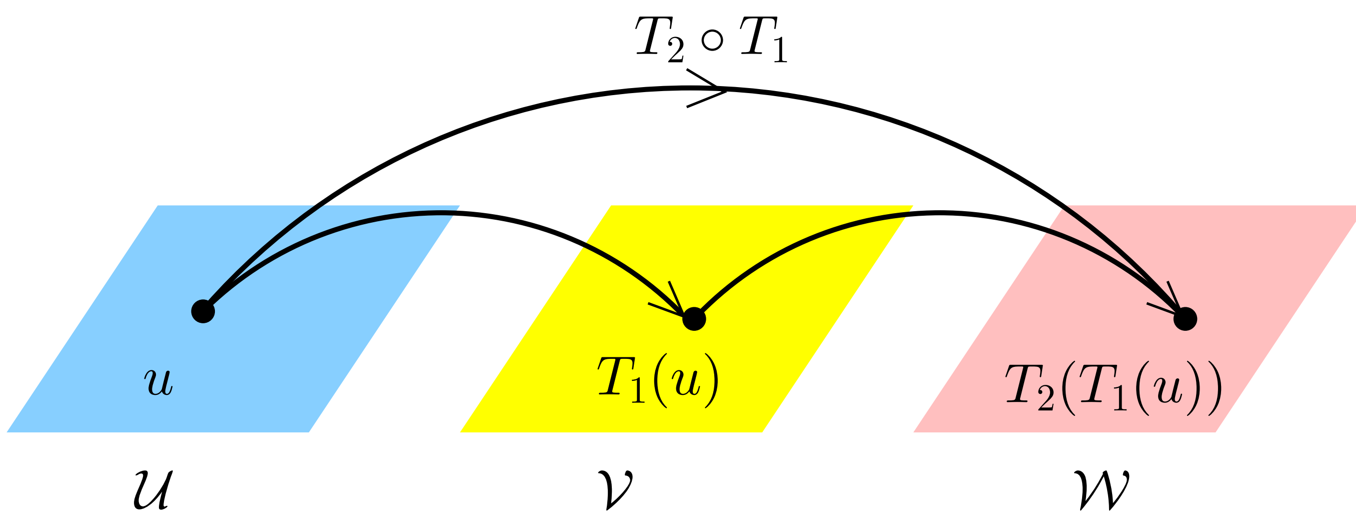
Composition of linear transformations

let $T_1 : \mathcal{U} \to \mathcal{V}$ and $T_2 : \mathcal{V} \to \mathcal{W}$ be linear transformations

the **composition** of T_2 *with* T_1 is the function defined by

$$(T_2 \circ T_1)(u) = T_2(T_1(u))$$

where u is a vector in $\mathcal{U}$



Theorem ✎ if T_1, T_2 are linear, so is $T_2 \circ T_1$

example 1: $T_1 : \mathbf{P}_1 \to \mathbf{P}_2, \quad T_2 : \mathbf{P}_2 \to \mathbf{P}_2$

$$T_1(p(t)) = tp(t), \quad T_2(p(t)) = p(2t+4)$$

then the composition of T_2 with T_1 is given by

$$(T_2 \circ T_1)(p(t)) = T_2(T_1(p(t))) = T_2(tp(t)) = (2t+4)p(2t+4)$$

example 2: $T : \mathcal{V} \to \mathcal{V}$ is a linear operator, $I : \mathcal{V} \to \mathcal{V}$ is identity operator

$$(T \circ I)(v) = T(I(v)) = T(v), \quad (I \circ T)(v) = I(T(v)) = T(v)$$

hence, $T \circ I = T$ and $I \circ T = T$

example 3: $T_1 : \mathbf{R}^n \to \mathbf{R}^m$, $T_2 : \mathbf{R}^m \to \mathbf{R}^n$ with

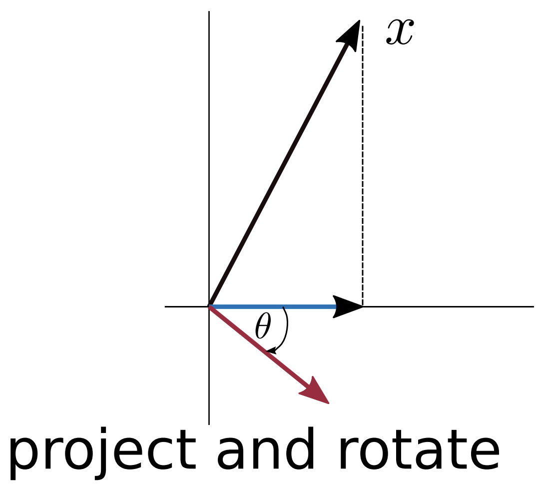
$$T_1(x) = Ax, \quad T_2(w) = Bw, \quad A \in \mathbf{R}^{m \times n}, B \in \mathbf{R}^{n \times m}$$

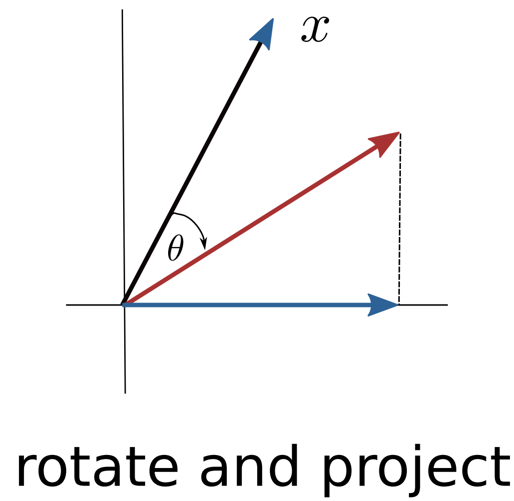
then $T_1 \circ T_2 = AB$ and $T_2 \circ T_1 = BA$

Order of operations matters

let $T_1, T_2 : \mathbf{R}^2 \to \mathbf{R}^2$ be the following matrix transformations

- $T_1(x)$ is the projection of x on the x_1-axis
- $T_2(x)$ is the rotation of x by θ (clockwise direction)





the composite of T_2 **with** T_1 VS the composite of T_1 **with** T_2

which is which ?

Nonlinear composite transformations

composite transformations can be defined for nonlinear mappings

many examples in applications:

- $T_1 : \mathbf{R}^n \to \mathbf{R}$ and $T_2 : \mathbf{R} \to \mathbf{R}$ **norm-squared**

$$T_1(x) = \|x\|_2, \quad T_2(x) = x^2 \quad \Rightarrow \quad (T_2 \circ T_1)(x) = \|x\|_2^2 = x^T x$$

- $T_1 : \mathbf{R}^n \to \mathbf{R}^n$ and $T_2 : \mathbf{R}^m \to \mathbf{R}$ **norm of affine**

$$T_1(x) = Ax + b, \quad T_2(x) = \|x\|_2^2 \quad \Rightarrow \quad (T_2 \circ T_1)(x) = \|Ax + b\|_2^2$$

- $T_1 : \mathbf{R}^n \to \mathbf{R}^m$ and $T_2 : \mathbf{R}^m \to \mathbf{R}^m$ **transform in neural network**

$$T_1(x) = Wx+b, \quad T_2(x) = \max(0, x) \;\Rightarrow\; (T_2 \circ T_1)(x) = \max(0, Wx+b)$$

Two operators cancel each other

scaling operators: $T_1, T_2 : \mathbf{R}^n \to \mathbf{R}^n$

$$\begin{aligned} T_1(x_1, x_2, \ldots, x_n) &= (a_1 x_1, a_2 x_2, \ldots, a_n x_n) \\ T_2(x_1, x_2, \ldots, x_n) &= (x_1/a_1, x_2/a_2, \ldots, x_n/a_n), \quad \forall a_k \neq 0 \\ (T_2 \circ T_1)(x) &= (T_1 \circ T_2)(x) = x \end{aligned}$$

shift operators: $T_1, T_2 : \mathbf{R}^n \to \mathbf{R}^n$

$$\begin{aligned} T_1(x_1, x_2, \ldots, x_n) &= (x_2, x_3, x_4, \ldots, x_n, x_1) \\ T_2(x_1, x_2, \ldots, x_n) &= (x_n, x_1, x_2, \ldots, x_{n-2}, x_{n-1}) \\ (T_2 \circ T_1)(x) &= T_2(x_2, x_3, \ldots, x_n, x_1) = x \\ (T_1 \circ T_2)(x) &= T_1(x_n, x_1, \ldots, x_{n-2}, x_{n-1}) = x \end{aligned}$$

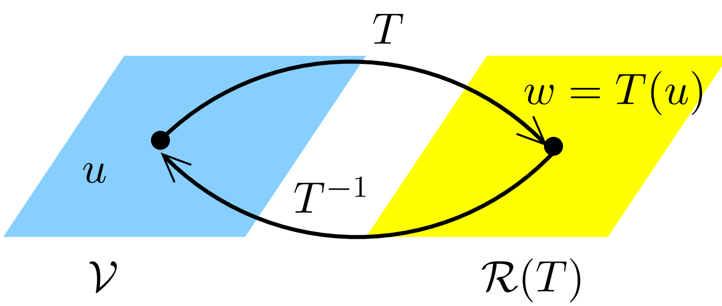
in these examples, T_2 brings the image under T_1 back to the original x !

Inverse of linear transformation

a linear transformation $T : \mathcal{V} \to \mathcal{W}$ is **invertible** if there is a transformation $S : \mathcal{W} \to \mathcal{V}$ satisfying

$$S \circ T = I_{\mathcal{V}} \quad \text{and} \quad T \circ S = I_{\mathcal{W}}$$

we call S the **inverse** of T and denote $S = T^{-1}$



$$T^{-1}(T(u)) = u \quad \forall u \in \mathcal{U}$$

$$T(T^{-1}(w)) = w \quad \forall w \in \mathcal{R}(T)$$

Facts:

- the inverse transformation $T^{-1} : \mathcal{R}(T) \to \mathcal{V}$ exists if and only if T is one-to-one
- $T^{-1} : \mathcal{R}(T) \to \mathcal{V}$ is also linear ✎

Inverse of matrix transformation

consider $T : \mathbf{R}^n \to \mathbf{R}^n$ where $T(x) = Ax$

- T is one-to-one if and only if A is invertible
- T^{-1} exists if and only if A is invertible

the inverse transformation must satisfy

$$T^{-1}(T(x)) = T^{-1}(Ax) = x, \quad \forall x \in \mathbf{R}^n$$

to find the description of T^{-1}

let $y = Ax$ and since A^{-1} exists, we can write $x = A^{-1}y$

$$T^{-1}(Ax) = T^{-1}(y) = A^{-1}y$$

this holds for all $y \in \mathbf{R}^n$ (since $y \in \mathcal{R}(A) = \mathbf{R}^n$)

conclusion: the inverse transformation is simply the matrix transformation given by A^{-1}

difference operator: $T : \mathbf{R}^n \to \mathbf{R}^n$

$$T(x) = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \\ \vdots \\ x_n - x_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} x \triangleq Ax$$

does T have an inverse ? if yes, what would it be ?

- please check ✎ that A is invertible and therefore T^{-1} exists
- $T^{-1}(x)$ is given

$$T^{-1}(x) = A^{-1}x = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ \vdots \\ x_1 + x_2 + \cdots + x_n \end{bmatrix}$$

T^{-1} is the cumulative sum operator ! (difference cancels with sum)

Inverse of transformation on $\mathbf{P}_n$

$T : \mathbf{P}_1 \to \mathbf{P}_1$, $T(p(x)) = p(x+c)$ where $c \in \mathbf{R}$ is given

- it can be verified ✎ that T is linear and one-to-one
- let $p(x) = a_0 + a_1 x$ be any polynomial in $\mathbf{P}_1$, T^{-1} must satisfy

$$T^{-1}(T(p(x)) = T^{-1}(a_0 + a_1(x+c)) = p(x) = a_0 + a_1 x, \quad \forall a_0, a_1 \in \mathbf{R}$$

- to find description of T^{-1}, let $q(x) = b_0 + b_1 x \triangleq a_0 + a_1(x+c)$ and we should write a_0, a_1 in terms of b_0, b_1

$$b_0 + b_1 x = a_0 + a_1 c + a_1 x \quad \Rightarrow \quad a_0 = b_0 - b_1 c, \;\; a_1 = b_1$$

- we can write $T^{-1}(b_0 + b_1 x) = b_0 - b_1 c + b_1 x = b_0 + b_1(x - c)$

it shows that $T^{-1}(q(x)) = q(x-c)$ (forward translation $x+c$ cancels with backward translation $x - c$)

Domain of T^{-1} may not be the whole co-domain of T

$T : \mathbf{R}^2 \to \mathbf{R}^{2\times 2}$ and given $a, c \neq 0$

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} ax_1 & 0 \\ 0 & cx_2 \end{bmatrix}$$

we can verify that ✎

- T is linear and one-to-one (hence, T^{-1} exists)
- $\mathcal{R}(T) = \operatorname{span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ (not the whole $\mathbf{R}^{2\times 2}$)

$T^{-1} : \mathcal{R}(T) \to \mathbf{R}^2$ is defined from $\mathcal{R}(T)$ and must satisfy

$$T^{-1}\left(\begin{bmatrix} ax_1 & 0 \\ 0 & cx_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

it follows that $T^{-1}(Y) = (y_{11}/a, y_{22}/c)$ where $Y \in \mathcal{R}(T)$

Composition of one-to-one linear transformation

if $T_1 : \mathcal{U} \to \mathcal{V}$ and $T_2 : \mathcal{V} \to \mathcal{W}$ are one-to-one linear transformation, then

- $T_2 \circ T_1$ is one-to-one
- $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$

example: $T_1 : \mathbf{R}^n \to \mathbf{R}^n, \quad T_2 : \mathbf{R}^n \to \mathbf{R}^n$

$$\begin{aligned} T_1(x_1, x_2, \ldots, x_n) &= (a_1x_1, a_2x_2, \ldots, a_nx_n), \quad a_k \neq 0, k = 1, \ldots, n \\ T_2(x_1, x_2, \ldots, x_n) &= (x_2, x_3, \ldots, x_n, x_1) \end{aligned}$$

both T_1 and T_2 are invertible and the inverses are

$$\begin{aligned} T_1^{-1}(w_1, w_2, \ldots, w_n) &= ((1/a_1)w_1, (1/a_2)w_2, \ldots, (1/a_n)w_n) \\ T_2^{-1}(w_1, w_2, \ldots, w_n) &= (w_n, w_1, \ldots, w_{n-1}) \end{aligned}$$

from a direct calculation, the composition of T_1^{-1} with T_2^{-1} is

$$\begin{aligned}(T_1^{-1} \circ T_2^{-1})(w) &= T_1^{-1}(w_n, w_1, \ldots, w_{n-1}) \\ &= ((1/a_1)w_n, (1/a_2)w_1, \ldots, (1/a_n w_{n-1}))\end{aligned}$$

now consider the composition of T_2 with T_1

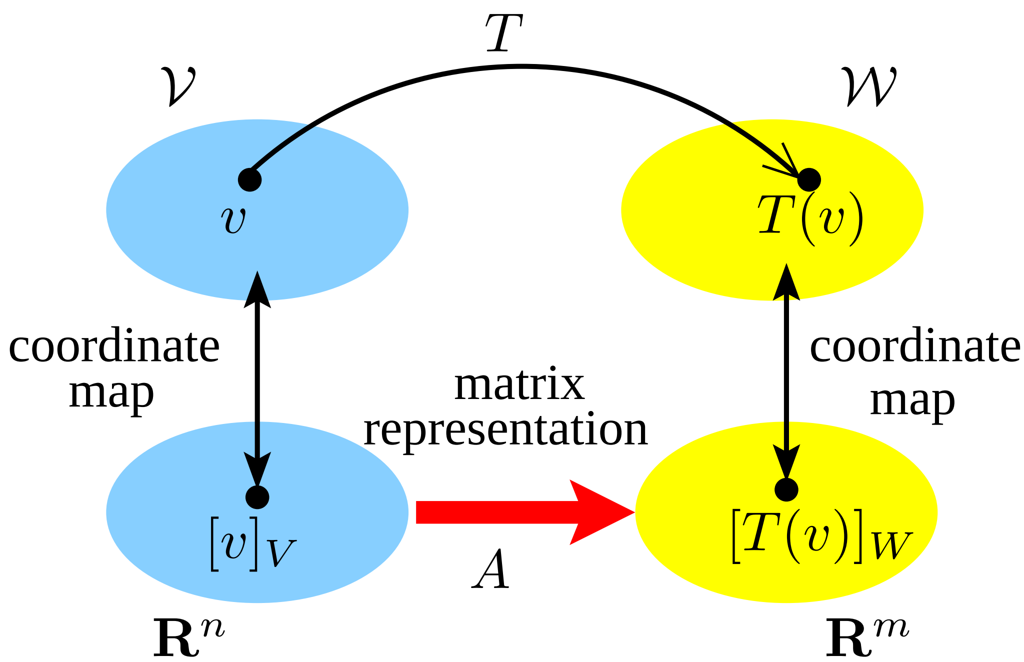
$$(T_2 \circ T_1)(x) = (a_2 x_2, \ldots, a_n x_n, a_1 x_1)$$

it is clear to see that

$$(T_2 \circ T_1) \circ (T_1^{-1} \circ T_2^{-1}) = I$$

Matrix representation for linear transformation

let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation



V is a basis for $\mathcal{V}$

$$\dim \mathcal{V} = n$$

W is a basis for $\mathcal{W}$

$$\dim \mathcal{W} = m$$

how to represent an image of T in terms of its coordinate vector ?

Problem: find a matrix $A \in \mathbf{R}^{m \times n}$ that maps $[v]_V$ into $[T(v)]_W$

key idea: the matrix A must satisfy

$$A[v]_V = [T(v)]_W, \quad \textbf{for all} \;\; v \in \mathcal{V}$$

hence, it suffices to hold *for all vector in a basis* for $\mathcal{V}$

suppose a basis for $\mathcal{V}$ is $V = \{v_1, v_2, \ldots, v_n\}$

$$A[v_1] = [T(v_1)], \quad A[v_2] = [T(v_2)], \quad \ldots, \quad A[v_n] = [T(v_n)]$$

(we have dropped the subscripts that refer to the choice of bases V, W

A is a matrix of size $m \times n$, so we can write A as

$$A = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}$$

where a_k's are the columns of A

the coordinate vectors of v_k's are simply the standard unit vectors

$$[v_1] = e_1, \quad [v_2] = e_2, \quad \ldots, \quad [v_n] = e_n$$

hence, we have

$$A[v_1] = a_1 = [T(v_1)], \quad A[v_2] = a_2 = [T(v_2)], \quad \cdots, \quad A[v_n] = a_n = [T(v_n)]$$

stack these vectors back in A

$$A = \begin{bmatrix} [T(v_1)] & [T(v_2)] & \cdots & [T(v_n)] \end{bmatrix}$$

- the columns of A are the coordinate maps of the images of the basis vectors in $\mathcal{V}$
- we call A the **matrix representation** for T relative to the bases V and W and denote it by
$$[T]_{W,V}$$
- a matrix representation *depends* on the **choice of bases** for $\mathcal{V}$ and $\mathcal{W}$

special case: $T : \mathbf{R}^n \to \mathbf{R}^m$, $T(x) = Bx$ we have $[T] = B$ relative to the *standard bases* for $\mathbf{R}^m$ and $\mathbf{R}^n$

example: $T : \mathcal{V} \to \mathcal{W}$ where

$$\begin{array}{rcl} \mathcal{V} & = & \mathbf{P}_1 \quad \text{with a basis} \quad V = \{1, t\} \\ \mathcal{W} & = & \mathbf{P}_1 \quad \text{with a basis} \quad W = \{t-1, t\} \end{array}$$

define $T(p(t)) = p(t+1)$, find $[T]$ relative to V and W

solution.

find the mappings of vectors in V and their coordinates relative to W

$$\begin{array}{rcccl} T(v_1) = T(1) & = & 1 & = & -1 \cdot (t-1) + 1 \cdot t \\ T(v_2) = T(t) & = & t+1 & = & -1 \cdot (t-1) + 2 \cdot t \end{array}$$

hence $[T(v_1)]_W = (-1, 1)$ and $[T(v_2)]_W = (-1, 2)$

$$[T]_{WV} = \begin{bmatrix} [T(v_1)]_W & [T(v_2)]_W \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}$$

example: given a matrix representation for $T : \mathbf{P}_2 \to \mathbf{R}^2$

$$[T] = \begin{bmatrix} 5 & 2 & -1 \\ 3 & 0 & 4 \end{bmatrix}$$

relative to the bases $V = \{2 - t, t + 1, t^2 - 1\}$ and $W = \{(1, 0), (1, 1)\}$

find the image of $6t^2$ under T

solution. find the coordinate of $6t^2$ relative to V by writing

$$6t^2 = \alpha_1 \cdot (2 - t) + \alpha_2 \cdot (t + 1) + \alpha_3 \cdot (t^2 - 1)$$

solving for $\alpha_1, \alpha_2, \alpha_3$ gives

$$[6t^2]_V = \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}$$

from the definition of $[T]$:

$$[T(6t^2)]_W = [T]_{WV}[6t^2]_V = \begin{bmatrix} 5 & 2 & -1 \\ 3 & 0 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 30 \end{bmatrix}$$

then we read from $[T(6t^2)]_W$ that

$$T(6t^2) = 8 \cdot (1,0) + 30 \cdot (1,1) = (38,30)$$

Matrix representation for linear operators

we say T is a **linear operator** if T is a linear transformation from $\mathcal{V}$ to $\mathcal{V}$

- typically we use the same basis for $\mathcal{V}$, says $V = \{v_1, v_2, \dots, v_n\}$
- a matrix representation for T relative to V is denoted by $[T]_V$ where

$$[T]_V = \left[\begin{array}{cccc} [T(v_1)] & [T(v_2)] & \dots & [T(v_n)] \end{array} \right]$$

Theorem ✌

- T is one-to-one if and only if $[T]_V$ is invertible
- $[T^{-1}]_V = ([T]_V)^{-1}$

what is the matrix (relative to a basis) for the identity operator ?

Matrix representation for composite transformation

if $T_1 : \mathcal{U} \to \mathcal{V}$ and $T_2 : \mathcal{V} \to \mathcal{W}$ are linear transformations

and U, V, W are bases for $\mathcal{U}, \mathcal{V}, \mathcal{W}$ respectively

then

$$[T_2 \circ T_1]_{W,U} = [T_2]_{W,V} \cdot [T_1]_{V,U}$$

example: $T_1 : \mathcal{U} \to \mathcal{V}$, $T_2 : \mathcal{V} \to \mathcal{W}$

$$\mathcal{U} = \mathbf{P}_1, \quad \mathcal{V} = \mathbf{P}_2, \quad \mathcal{W} = \mathbf{P}_3$$

$$U = \{1, t\}, \quad V = \{1, t, t^2\}, \quad W = \{1, t, t^2, t^3\}$$

$$T_1(p(t)) = T_1(a_0 + a_1 t) = 2a_0 - 3a_1 t$$

$$T_2(p(t)) = 3tp(t)$$

find $[T_2 \circ T_1]$

solution. first find $[T_1]$ and $[T_2]$

$$\begin{aligned} T_1(1) &= 2 &&= 2\cdot 1 + 0\cdot t + 0\cdot t^2 \\ T_1(t) &= -3t &&= 0\cdot 1 - 3\cdot t + 0\cdot t^2 \end{aligned} \quad\Longrightarrow\quad [T_1] = \begin{bmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \end{bmatrix}$$

$$\begin{aligned} T_2(1) &= 3t &&= 0\cdot 1 + 3\cdot 1 + 0\cdot t^2 + 0\cdot t^3 \\ T_2(t) &= 3t^2 &&= 0\cdot 1 + 0\cdot 1 + 3\cdot t^2 + 0\cdot t^3 \\ T_2(t^2) &= 3t^3 &&= 0\cdot 1 + 0\cdot 1 + 0\cdot t^2 + 3\cdot t^3 \end{aligned} \quad\Longrightarrow\quad [T_2] = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

next find $[T_2 \circ T_1]$

$$\begin{aligned} (T_2\circ T_1)(1) &= T_2(2) &&= 6t \\ (T_2\circ T_1)(t) &= T_2(-3t) &&= -9t^2 \end{aligned} \quad\Longrightarrow\quad [T_2\circ T_1] = \begin{bmatrix} 0 & 0 \\ 6 & 0 \\ 0 & -9 \\ 0 & 0 \end{bmatrix}$$

easy to verify that $[T_2 \circ T_1] = [T_2]\cdot[T_1]$

References

Chapter 8 in

H. Anton, *Elementary Linear Algebra*, 10th edition, Wiley, 2010