Linear Transformation

Linear Algebra

Department of Computer Engineering

Sharif University of Technology

Hamid R. Rabiee rabiee@sharif.edu

Maryam Ramezani maryam.ramezani@sharif.edu

Overview

- Introduction
- Linear Transformation (Linear Map)
- Rotation-Projection-Reflection
- Non−linear Maps
- Null space and Range
- Onto and One-to-One Linear Transformation
- Fundamental Theorem of Linear Maps

Introduction

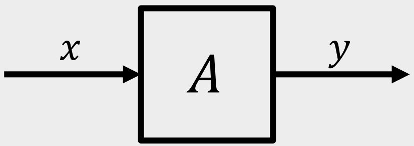
Linear Transformation

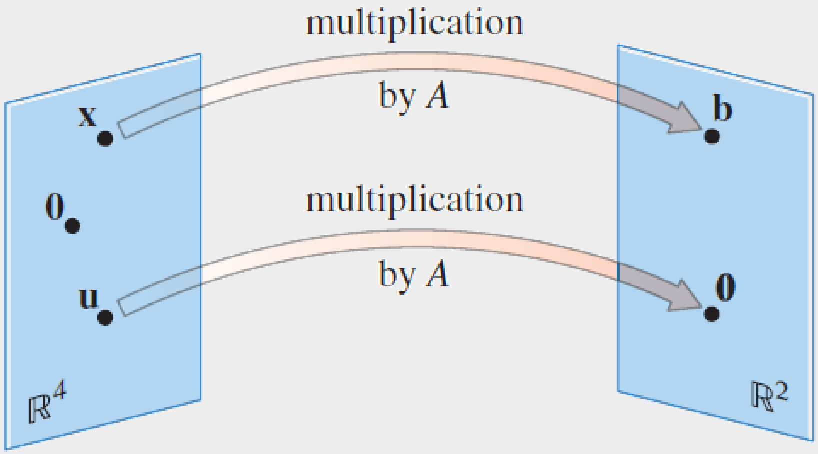
- Matrix is a linear transformation: map one vector to another vector

$$A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^{n}, y \in \mathbb{R}^{m}: \qquad y_{m\times 1} = A_{m\times n}x_{n\times 1}$$

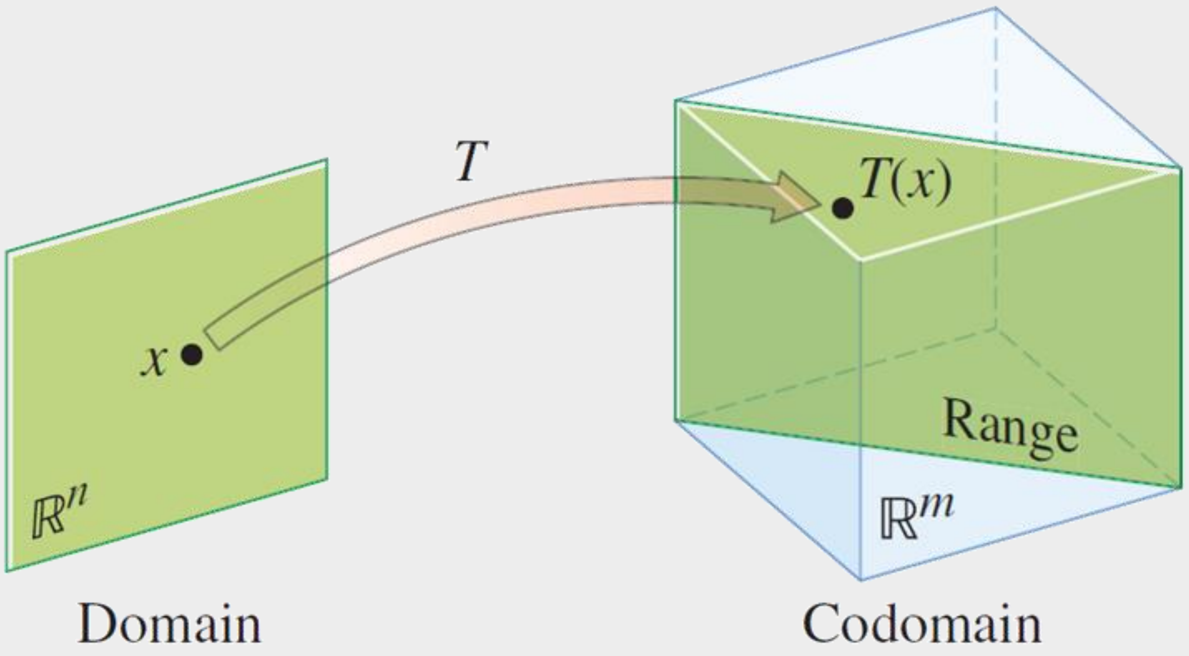
$$A : \mathbb{R}^{n} \to \mathbb{R}^{m}$$

- Input–output





Linear Transformation



Domain, codomain, and range of $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$

Example

Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, and define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$, so that

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

a. Find $T(\mathbf{u})$, the image of $\mathbf{u}$ under the transformation T.

b. Find an $\mathbf{x}$ in $\mathbb{R}^2$ whose image under T is $\mathbf{b}$.

c. Is there more than one $\mathbf{x}$ whose image under T is $\mathbf{b}$?

d. Determine if $\mathbf{c}$ is in the range of the transformation T.

Linear Transformation (Linear Map)

Linear mapping

Definition

Let V and W be vector spaces over the field $\mathbb{F}$. A linear transformation (or a linear map) from V into W is a function $\mathbf{T}: V \rightarrow W$ that satisfies following properties for all x, y in V and all scalars a in $\mathbb{F}$:

$$T(x+y) = T(x) + T(y)$$
$$T(\alpha x) = \alpha T(x)$$

Notes

- ❑ $T(0) = 0$
- ❑ Transformation preserves linear combinations

$$T(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1 \big(\mathrm{T}(\mathrm{x}_1)\big) + \cdots + \alpha_n \big(\mathrm{T}(\mathrm{x}_\mathrm{n})\big)$$

Linear mapping

Notes

- The set of linear maps from V to W is denoted by $\mathcal{L}(\mathrm{V}, \mathrm{W})$.
- The set of linear maps from V to V is denoted by $\mathcal{L}(\mathrm{V})$.
  In other words, $\mathcal{L}(\mathrm{V}) = \mathcal{L}(\mathrm{V}, V)$

Linear mapping

Theorem

Let $(v_1, \dots, v_n)$ be a ordered basis of finite-dimensional vector space V over the field $\mathbb{F}$ and $(w_1, \dots, w_n)$ an arbitrary list of any vectors in W.

Then there exists a unique linear map

$$T : V \to W \qquad \text{such that } T(v_i) = w_i.$$

Proof

Example

Which are linear mapping?

- ❑ **zero** map $0 : V \to W$
- ❑ **identity** map $I : V \to V$
- ❑ Let $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ be the **differentiation** map defined as $T_{\mathcal{P}(z)} = \mathcal{P}\acute{}(z)$
- ❑ Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the map given by $T(x, y) = (x - 2y, 3x + y)$
- ❑ $T(x) = e^x$

$T(x_1, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n)$

- ❑ $T : \mathbb{F} \to \mathbb{F}$ given by $T(x) = x - 1$

Algebraic Operations on L(V,W)

Definition

Let S and $T \in L(V, W)$ and $\lambda \in \mathbb{F}$. The sum $S + T$ and the product λT are the linear maps from V to W defined by:

$$(S + T)(v) = Sv + Tv \text{ and } (\lambda T)(v) = \lambda(Tv)$$

For all $v \in V$.

Theorem

With the addition and scalar multiplication as defined above, $L(V, W)$ is a vector space.

Proof

Review: Vector Space Properties

- **Addition of vector space** $(x + y)$
  - **Commutative** $x + y = y + x \ \forall x, y \in V$
  - **Associative** $(x + y) + z = x + (y + z) \ \forall x, y, z \in V$
  - **Additive identity** $\exists \mathbf{0} \in V$ such that $x + \mathbf{0} = x, \forall x \in V$
  - **Additive inverse** $\exists (-x) \in V$ such that $x + (-x) = 0, \forall x \in V$

- **Action of the scalars field on the vector space** (αx)

  - **Associative** $\alpha(\beta x) = (\alpha\beta)x \qquad \forall \alpha, \beta \in F; \forall x \in V$

  - **Distributive over** ……

    scalar addition: $(\alpha + \beta)x = \alpha x + \beta x \qquad \forall \alpha, \beta \in F; \forall x \in V$

    vector addition: $\alpha(x + y) = \alpha x + \alpha y \qquad \forall \alpha \in F; \forall x, y \in V$

  - **Scalar identity** $1x = x \qquad \forall x \in V$

Product of Linear maps

Definition

Let $T \in L(U, V)$ and $S \in L(V, W)$, then the product $ST \in L(U, W)$ is defined by:

$$(ST)(u) = S(Tu)$$

For all $u \in U$.

Notes

Note that ST is defined only when T maps into the domain of S. You should verify that ST is indeed a linear map from U to W whenever $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$.

Product of Linear maps

Notes

Multiplication of linear maps is not commutative.

Example

$$D \in L(P(R)) \;\; as \; D(P(x)) = P'(x)$$

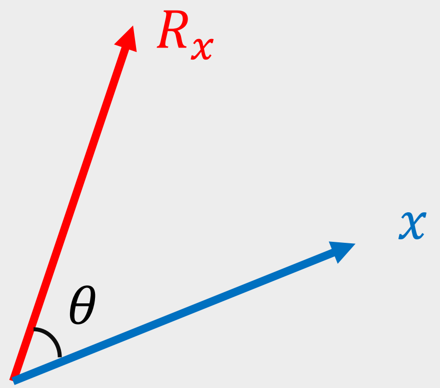
$$T \in L(P(R)) \;\; as \; T(P(x)) = x^2 P(x)$$

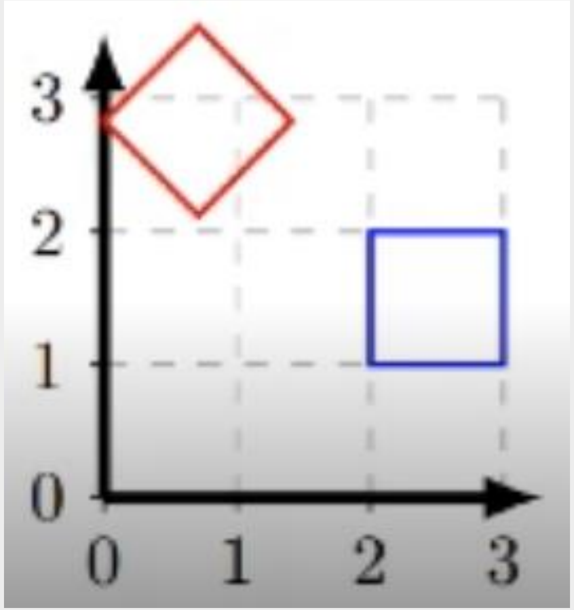
$$TD \neq DT$$

Rotation–Projection–Reflection

Rotation with $\boldsymbol{\theta}$ degree

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

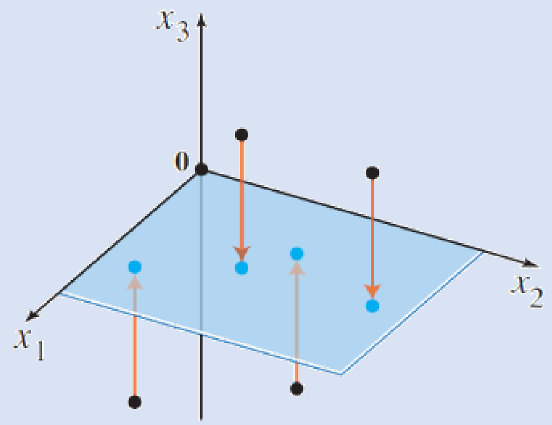




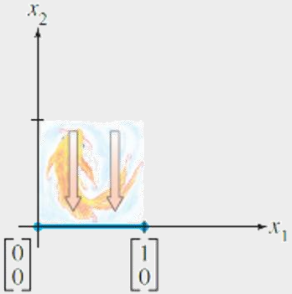
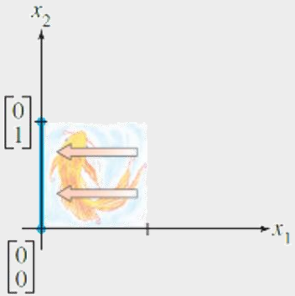
Example

If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ projects points in $\mathbb{R}^3$ onto the x_1x_2–plane because

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$



Projection

Transformation	Image of the Unit Square	Standard Matrix
Projection onto the x_1–axis		$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Projection onto the x_2–axis		$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Projection

Theorem

Suppose that V is a vector space and $P : V \to V$ is a linear transformation.

a) If $P^2 = P$ then P is called a **projection**.

b) If V is an inner product space and $P^2 = P = P^*$ then P is called an **orthogonal projection**.

We furthermore say that P **projects onto** $\mathrm{range}(P)$.

- Projection of vector v on:
  - Two orthogonal vectors
  - Two non-orthogonal vectors

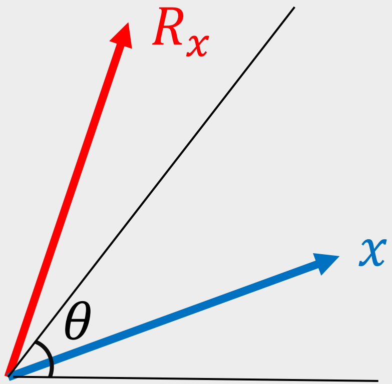
Projection on $\boldsymbol{\theta}$ Line

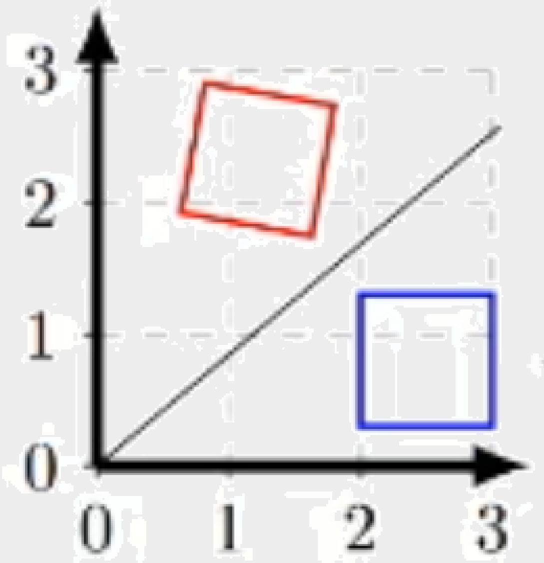
$$P = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

$$P^2 = P$$

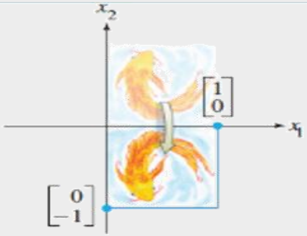
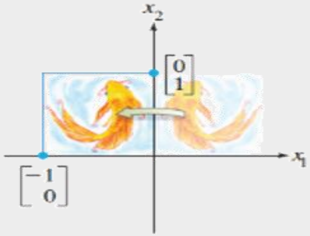
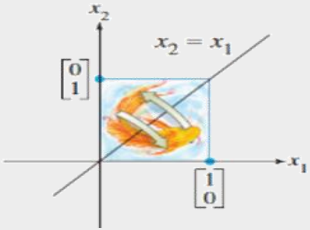
Reflection in the θ Line

- $R = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}$



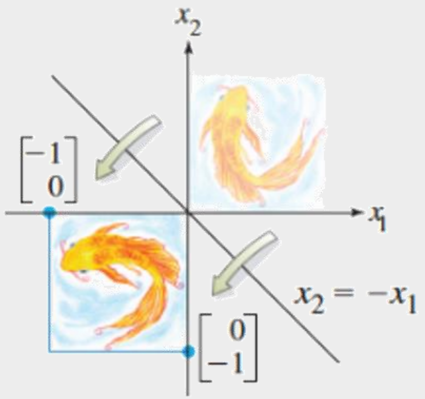


$$R^2 = I$$

Transformation	Image of the Unit Square	Standard Matrix
Reflection through the x_1–axis		$\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$
Reflection through the x_2–axis		$\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}$
Reflection through the line $x_2 = x_1$		$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$

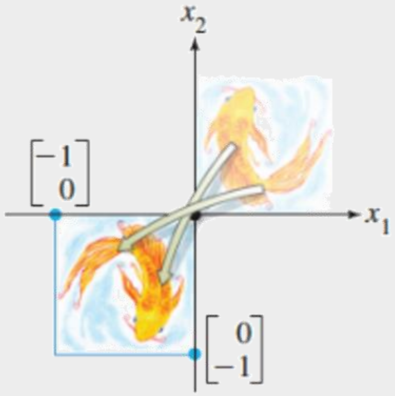
Reflection

Reflection through the line $x_2 = -x_1$



$$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

Reflection through the origin

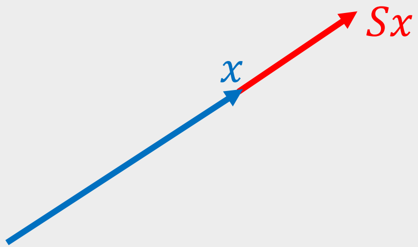


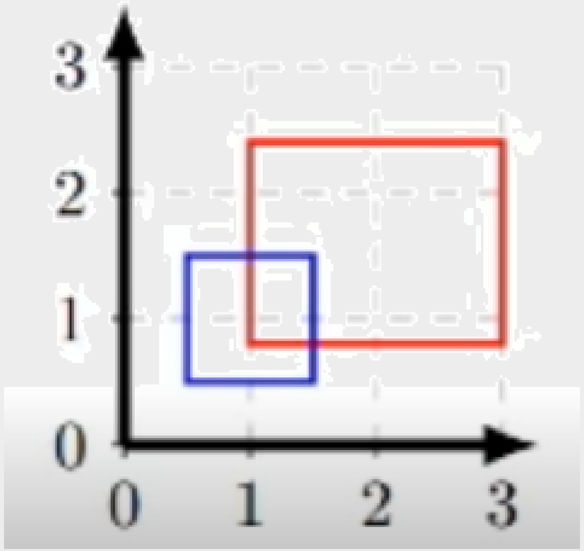
$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Applications

Uniform Scaling

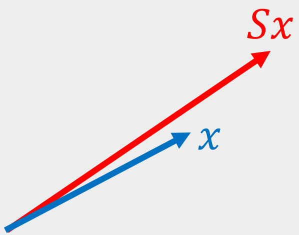
- $S = sI = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}$

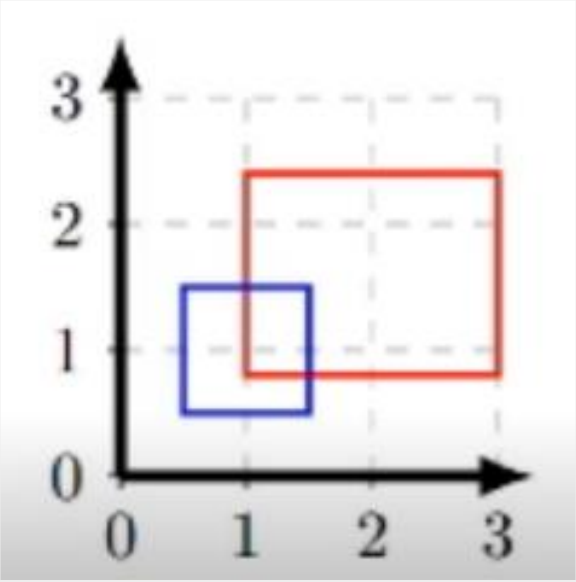




Non-uniform Scaling

- $S = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$





Shearing

Example

Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. The transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$

A typical shear matrix is of the form

$$S = \begin{pmatrix} 1 & 0 & 0 & \lambda & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

sheep

sheared sheep

A shear parallel to the x axis results in $\acute{x} = x + \lambda y$ and $\acute{y} = y$.

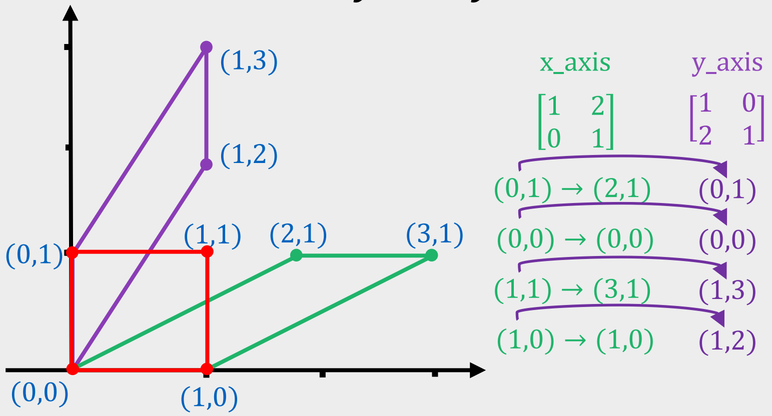
In matrix form:

$$\begin{pmatrix} \acute{x} \\ \acute{y} \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Similarly, a shear parallel to the y axis has $\acute{x} = x$ and $\acute{y} = y + \lambda x$.

In matrix form:

$$\begin{pmatrix} \acute{x} \\ \acute{y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$



Difference Matrix

Note

$$D_{(n-1)\times n} = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}$$

$$D: \mathbb{R}^n \to \mathbb{R}^{n-1} \quad \Rightarrow \quad D\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ \vdots \\ x_n - x_{n-1} \end{bmatrix}$$

Example

$$\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ -1 \\ 3 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} -1-0 \\ 3-(-1) \\ 2-3 \\ 5-2 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ -1 \\ 3 \end{bmatrix}$$

Selectors

- an $m \times n$ *selector matrix*: each row is a unit vector (transposed)

$$A = \begin{bmatrix} e_{k_1}^T \\ \vdots \\ e_{k_m}^T \end{bmatrix}$$

multiplying by A selects entries of x:

$$Ax = (x_{k_1}, x_{k_2}, \ldots, x_{k_m})$$

- $A : \mathbb{R}^n \to \mathbb{R}^m \quad \Rightarrow \quad A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{k_1} \\ x_{k_2} \\ \vdots \\ x_{k_m} \end{bmatrix}$

Example

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

- ❑ Selecting first and last elements of vector:
- ❑ Reversing the elements of vector:

Slicing

- Keeping m elements from r to s (m=s−r+1)

$$\begin{bmatrix} 0_{m\times(r-1)} & I_{m\times m} & 0_{m\times(n-s)} \end{bmatrix}$$

Example

- Slicing two first and one last elements:

$$\begin{bmatrix} -1 \\ 2 \\ 0 \\ -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$$

Down Sampling

- Down sampling with k: selecting one sample in every k samples

Example

$\mathrm{K} = 2?$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ \vdots \end{bmatrix} = \begin{bmatrix} x_1 \\ x_3 \\ x_5 \\ \vdots \end{bmatrix}$$

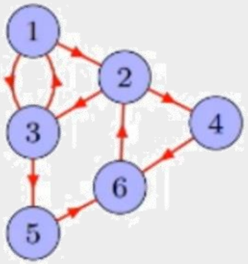
(i) $\sin 2A = 2 \sin A \cos A$

(ii) $\cos 2A = \cos^2 A - \sin^2 A$

- Rotation matrix

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \Rightarrow R^n = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^n = \begin{bmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{bmatrix}$$

- Adjacency matrix



$$A = \begin{array}{c} \begin{array}{cccccc} n1 & n2 & n3 & n4 & n5 & n6 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

$$A^2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Non-Linear Map

Norms

- First, the triangle inequality defines: $\| x + y \| \leq \| x \| + \| y \|$. Whereas the first requirement for linear mappings demands: $T(x + y) = T(x) + T(y)$. The problem here is in the $\leq$ condition, which means adding two vectors and then taking the norm can be less than the sum of the norms of the individual vectors. Such condition is, by definition, not allowed for linear mappings.

- Second, the positive definite defines: $// x // \geq 0$ and $// x // = 0 \Leftrightarrow x = 0$. Put simply, norms have to be a positive value. For instance, the norm of $// -x // = // x //$, instead of $// -x //$. But, the second property for linear mappings requires $// -\alpha x // = -\alpha \; // x //$. Hence, it fails when we multiply by a negative number (i.e., it can preserve the negative sign).

Translation

- Translation is a geometric transformation that moves every vector in a vector space by the same distance in a given direction. Translation is an operation that matches our everyday life intuitions: move a cup of coffee from your left to your right, and you would have performed translation in R3 space.
- $T: R^2 \to R^3$

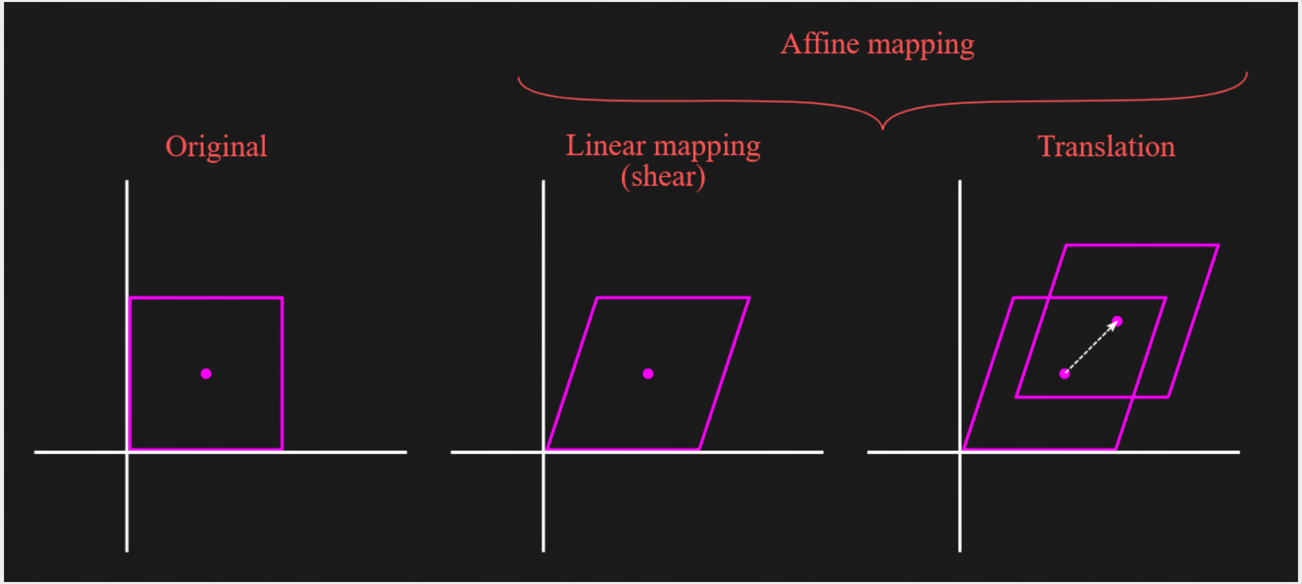
$$T_v = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$$

Affine mappings

- linear mapping + translation

$$M(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

where A is a linear mapping or transformation and $\mathbf{b}$ is the translation vector.



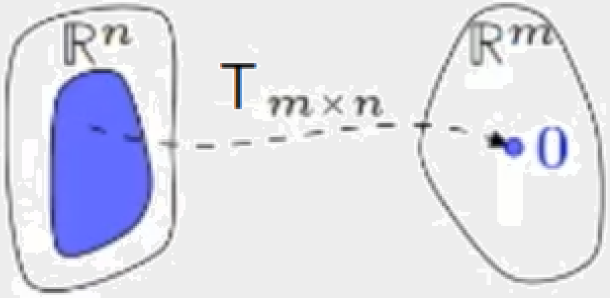
Null Spaces and Ranges

Definition

Let $T: V \to W$ be a linear map. Then the **null space or kernel of T** is the set of all vectors in V that map to zero:

$$N(T) = Null(T) = \{v \in V \mid Tv = 0\}$$

- $Nullity(T) := Dim(Null(T))$



Null Space

Theorem

Suppose $T \in L(V, W)$. Then null T is a subspace of V.

Proof

Theorem

Suppose $T \in L(V, W)$. Then null T is vector space.

Example

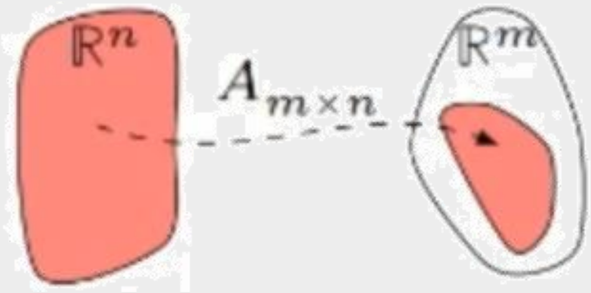
Find Null Space T?

- ❑ **zero** map $0 : V \to W$
- ❑ Let $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ be the **differentiation** map defined as $T_{\mathcal{P}(z)} = \acute{\mathcal{P}}(z)$
- ❑ Let $T : C^3 \to C$ be the map given by $T(x, y, z) = x + 2y + 3z$
- ❑ $T(P(x)) = x^2 P(x)$
- ❑ $T \in L(\mathbb{F}^\infty)$ given by $T(x_1, x_2, \dots) = T(x_2, x_3, \dots)$
- ❑ When is Nullity(T) = 0 ?

Definition

Let $T: V \to W$ be a linear map. Then the **range of T** is the subset of W consisting of those vectors that are equal to Tv for some $v \in V$:

$$range(T) = \{T(v) | v \in V\}$$



Theorem

Suppose $T \in L(V, W)$. Then range T is a subspace of V.

Proof

Theorem

Suppose $T \in L(V, W)$. Then range T is vector space.

Example

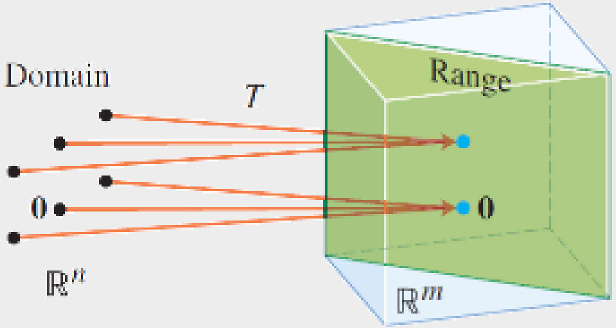
Find Range T?

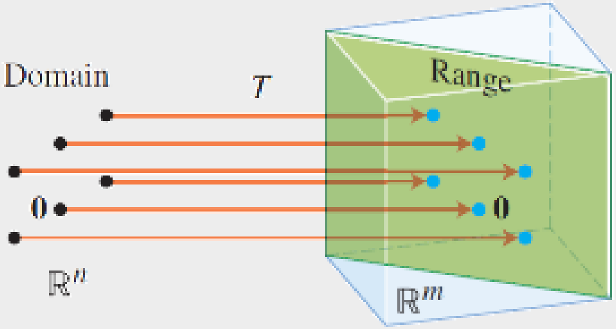
- **zero** map $0 : V \to W$
- Let $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ be the **differentiation** map defined as $T_{\mathcal{P}(z)} = \acute{\mathcal{P}}(z)$

One-to-one (Injective)

One-to-One Mapping

- A mapping T : $\mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be one-to-one (injective) $\mathbb{R}^m$ if each **b** in $\mathbb{R}^m$ is the image of *at most one* **x** in $\mathbb{R}^n$





Injective and homogeneous linear equation

Theorem

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if the equation T(x) = 0 has only the trivial solution.

Proof

One-to-One and Null Space

Theorem

Let $T: V \rightarrow W$ be a linear transformation. Then T is one-to-one if and only if the equation Null(T)={0} (Nullity(T)=0!).

Proof

Example

Let T be the linear transformation whose standard matrix is

$$A = \begin{pmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

Does T map $\mathbb{R}^4$ onto $\mathbb{R}^3$? Is T a one-to-one mapping?

One-to-One Linear Transformation

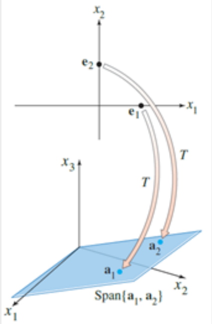
Important

Let $\mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let A be the standard matrix for T. Then:

a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$.
b. T is one-to-one if and only if the columns of A are linearly independence.

Example

Let $T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$. Show that T is a one-to-one linear transformation. Does T map $\mathbb{R}^2$ onto $\mathbb{R}^3$?



One-to-One Transformations

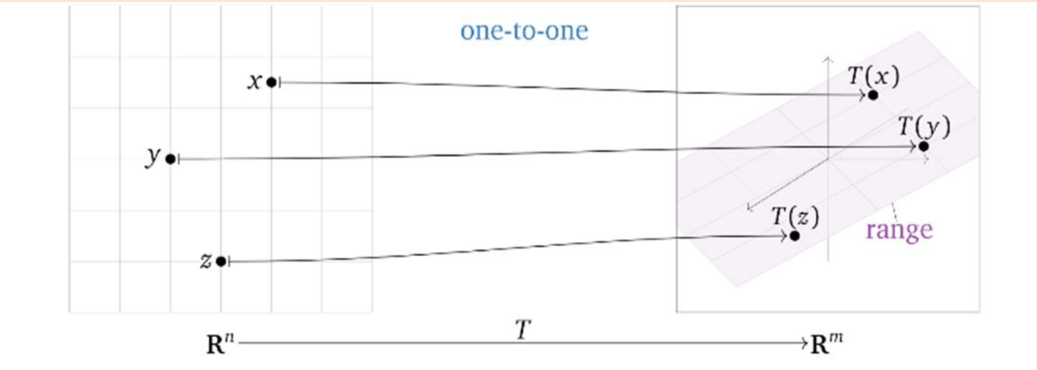
Definition

One-to-one transformations: A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if, for every vector b in $\mathbb{R}^m$, the equation T(x) = b has at most one solution x in $\mathbb{R}^n$.

Remark

Here are some equivalent ways of saying that T is one-to-one:

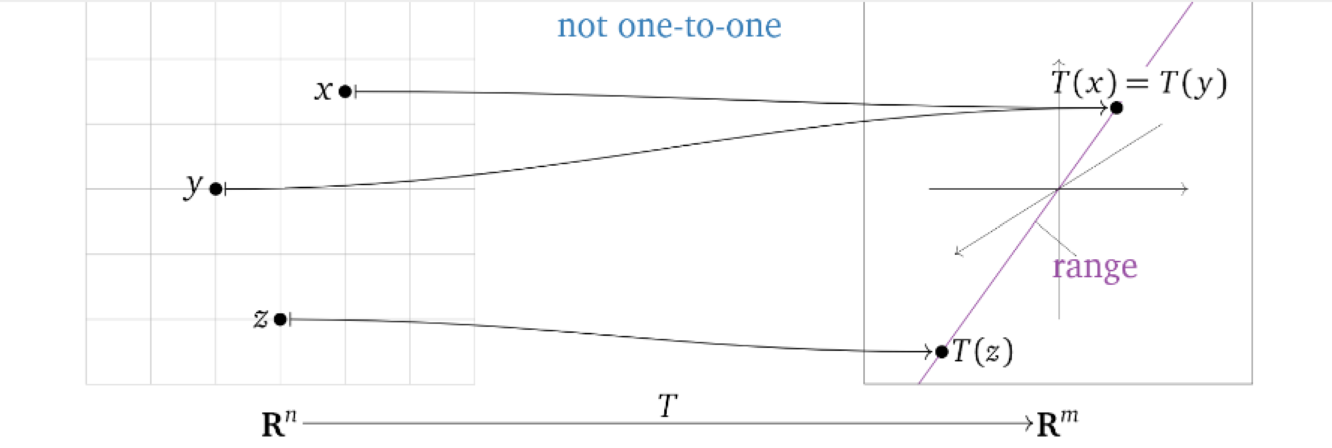
- For every vector b in $\mathbb{R}^m$, the equation T(x) = b has zero or one solution x in $\mathbb{R}^n$.
- Different inputs of T have different outputs.
- If T(u) = T(v) then u = v.



Remark

Here are some equivalent ways of saying that T is not one-to-one:

- There exist some vector b in $\mathbb{R}^m$ such that the equation T(x) = b has more than one solution x in $\mathbb{R}^n$.
- There are two different inputs of T with the same output.
- There exist vectors u, v such that $u \neq v$ but T(u) = T(v).



One-to-one Transformations

Theorem

Let A be an $m \times n$ matrix and let $T(x) = Ax$ be the associated matrix transformation. The following statements are equivalent:

1. T is one-to-one.
2. For every b in $\mathbb{R}^m$, the equation $T(x) = b$ has at most one solution.
3. For every b in $\mathbb{R}^m$, the equation $T(x) = b$ has a unique solution or is inconsistent.
4. $Ax = 0$ has only the trivial solution.
5. The columns of A are linearly independent.
6. A has a pivot in every column.
7. The range of T has dimension n.

One-to-one Transformations

Important

Wide matrices do not have one-to-one transformations.

If $T: \mathbb{R}^n \to \mathbb{R}^m$ is an one-to-one matrix transformation, what can we say about the relative sizes of n and m?

The matrix associated to T has n columns and m rows. Each row and each column can only contain one pivot, so in order for A to have a pivot in every column, it must have at least as many rows as columns:

$n \leq m.$

This says that for instance, $\mathbb{R}^3$ is **too big** to admit a one-to-one linear transformation into $\mathbb{R}^2$.

Note that there exist tall matrices that are not one-to-one, for example,

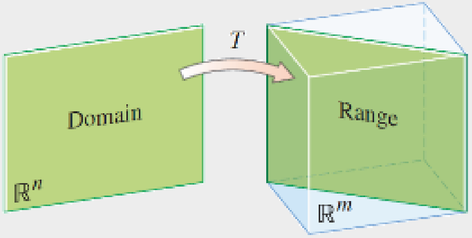
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Does not have a pivot in every column.

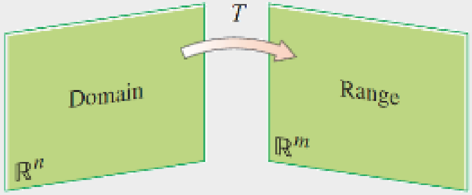
Onto (Surjective) Linear Transformation

Onto Mapping

- A mapping T : $\mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be **onto (surjective)** $\mathbb{R}^m$ if each **b** in $\mathbb{R}^m$ is the image of *at least one* **x** in R^n



T is *not* onto $\mathbb{R}^m$



T is onto $\mathbb{R}^m$

Onto (surjective) Transformation

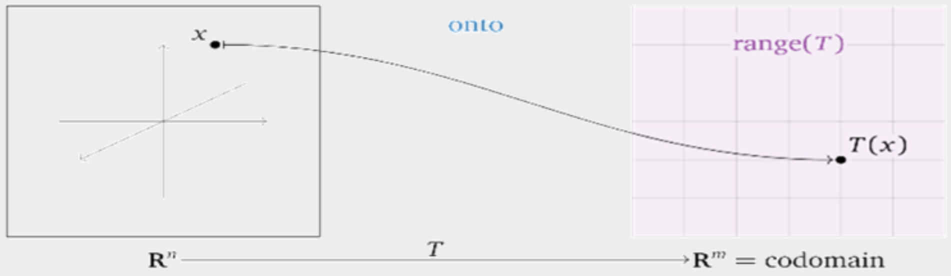
Definition

A transformation $T : V \to W$ is onto if, for every vector b in W, the equation T(x) = b has at least one solution x in V. It range equals W.

Note

Here are some equivalent ways of saying that T is onto:

- The range of T is equal to the codomain of T.
- Every vector in the codomain is the output of some input vector.



Onto Transformations

Example

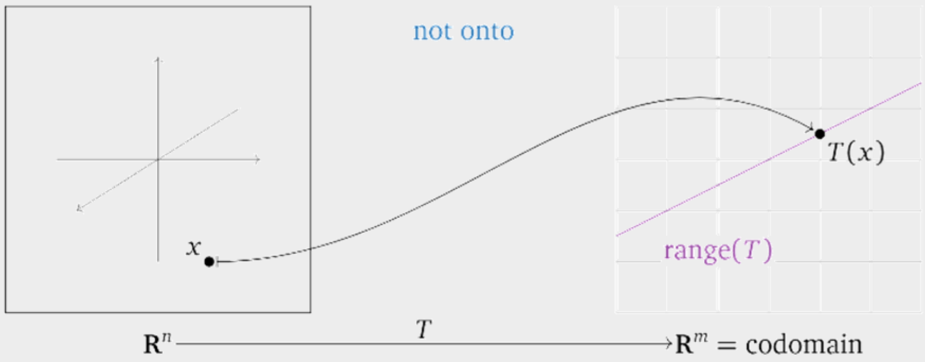
Which one is surjective?

- ❑ $D \in L\big(P_5(R)\big)$ defined by $DP = P'$
- ❑ $S \in L(P_5(R), P_4(R))$ defined by $SP = P'$

Note

Here are some equivalent ways of saying that T is not onto:

- The range of T is smaller to the codomain of T.
- There exists a vector b in $\mathbb{R}^m$ such that the equation T(x) = b does not have a solution
- There is a vector in the codomain that is not the output of any input vector.



Onto Transformation

Theorem

Let A be an $m \times n$ matrix and let T(x) = Ax be the associated matrix transformation. The following statement are equivalent:

- T in onto.
- T(x) = b has at least one solution for every b in $\mathbb{R}^m$.
- Ax = b is consistent for every b in $\mathbb{R}^m$.
- The columns of A span $\mathbb{R}^m$.
- A has a pivot in every row.
- The range of T has dimension m.

Onto Transformations

Important

Tall matrices do not have onto transformations.

If $T: \mathbb{R}^n \to \mathbb{R}^m$ is an onto matrix transformation, what can we say about the relative sizes of n and m?

The matrix associated to T has n columns and m rows. Each row and each column can only contain one pivot, so in order for A to have a pivot in every row, it must have at least as many columns as rows: $m \leq n$.

This says that for instance, $\mathbb{R}^2$ is **too small** to admit an onto linear transformation to $\mathbb{R}^3$.

Note that there exist wide matrices that are not onto, for example,

$$\begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{pmatrix}$$

Does not have a pivot in every row.

Solution

The reduction row echelon form of A is :

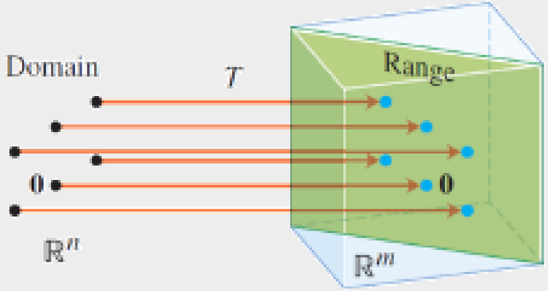
$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

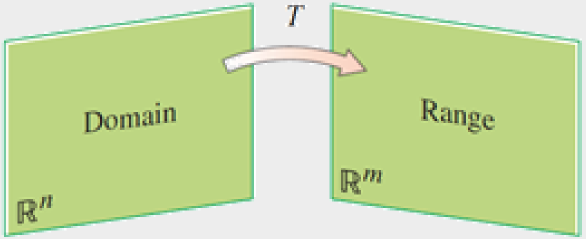
There is not a pivot in every row, so T is not onto. The range of T is the column space of A which is equal to

$$span\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \end{pmatrix}\right\} = span\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\}$$

since all three columns of A are collinear. Therefore, any vector not on the line through $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ is not in the range of T. for instance, if b = $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ then T(x) = b has no solution.

A is an m × n matrix, and T: $\mathbb{R}^n \to \mathbb{R}^m$ is the matrix transformation T(x) = Ax.





T is one-to-one	T is onto
$T(x) = b$ has *at most one solution* for every b.	$T(x) = b$ has *at least one solution* for every b.
The columns of A are linearly independent.	The columns of A span $\mathbf{R}^m$.
A has a pivot in every column.	A has a pivot in every row.
The range of T has dimension n.	The range of T has dimension m.

One-to-one and onto

Important

One-to-one is the same as onto for square matrices. We observed that a square has a pivot in every row if and only if it has a pivot in every column. Therefore, a matrix transformation T from $\mathbb{R}^n$ to itself is one-to-one if and only if it is onto : in this case, the two notations are equivalent.

Conversely, by this note, if a matrix transformation T: $\mathbb{R}^m \to \mathbb{R}^n$ is both one-to-one and onto, then
m = n.

Note that in general, a transformation T is both one-to-one and onto if and only if T(x) = b has exactly one solution for all b in $\mathbb{R}^m$.

Bijective

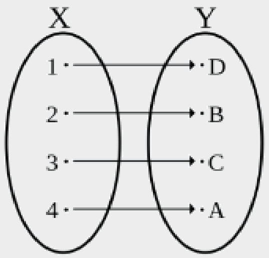
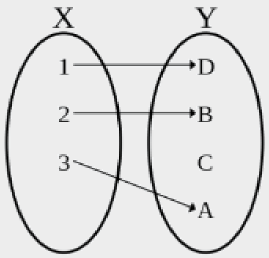
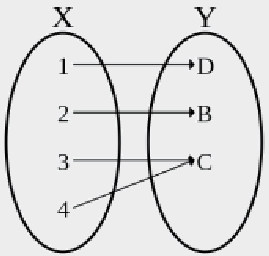
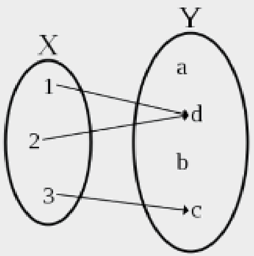
Note

- One-to-one and onto.
- If and only if every possible image is mapped to by exactly one argument.

Conclusion

onto

One-to-one

	surjective	non-surjective
injective	 bijective	 injective-only
non-injective	 surjective-only	 general

Machine learning application

- The central problem in machine learning and deep learning is to meaningfully transform data; in other words, to learn useful representations of the input data at hand – representations that get us to the expected output.

Fundamental Theorem of Linear Maps

dim V = dim null T + dim range

Theorem

Let V be a finite-dimensional vector space and $T \in L(V, W)$. Then rang T is finite-dimensional and

$$\mathrm{Dim}(V) = Nullity(T) + Dim(range(T))$$

Proof

Corollary

Linear map to a lower-dimensional space is not injective.

Proof

Corollary

Linear map to a higher-dimensional space is not surjective

Proof

Example

Is T injective or not?

$$T\colon \mathbb{F}^4 \to F^3$$

$$T(x_1, x_2, x_3, x_4) = (\sqrt{7}x_1 + \pi x_2 + x_4, 97x_1 + 3x_2 + 2x_3, x_2 + 6x_3 + 7x_4)$$

References

- Chapter 1: Advanced Linear and Matrix Algebra, Nathaniel Johnston
- Chapter 6: Linear Algebra David Cherney
- Linear Algebra and Optimization for Machine Learning
- Introduction to Applied Linear Algebra Vectors, Matrices, and Least Squares