Lecture 11, Oct 3, 2025

Linear Maps and Matrix Representations

Definition

A function $f : \mathcal{X} \mapsto \mathcal{Y}$ is *injective* (one-to-one) if

$$\forall x_1, x_2 \in \mathcal{X},\ f(x_1) = f(x_2) \implies x_1 = x_2$$

or contrapositively $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$, i.e. different inputs always map to different outputs.
f is *surjective* (onto) if

$$\forall y \in \mathcal{Y}, \exists x \in \mathcal{X} \text{ s.t. } f(x) = y$$

i.e. the output reaches the entirety of $\mathcal{Y}$.

A function that is both injective and surjective is called *bijective*.

Definition

Let $\mathcal{X}, \mathcal{Y}$ be vector spaces, then a function $L : \mathcal{X} \mapsto \mathcal{Y}$ is a *linear transformation* (or *linear map*) if

$$\forall x_1, x_2 \in \mathcal{X}, \lambda \in \mathbb{F},\ L(x + \lambda y) = L(x) + \lambda L(y)$$

- Consider finite dimensional vector spaces $\mathcal{X}, \mathcal{Y}$ where $\{ x^1, \dots, x^n \}$ is a basis for $\mathcal{X}$ and $\{ y^1, \dots, y^m \}$ is a basis for $\mathcal{Y}$
  - For each x_i, $L(x_i) \in \mathcal{Y}$ so it can be expressed as coordinates $L(x^i) = \sum_{j=1}^{m} a_{ji} y^j$
  - From this, we can form $\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ where column i contains the coordinates of $L(x^i)$
  - Now consider $x \in \mathcal{X} \implies x = \sum_{i=1}^{n} c_i x^i$ and $y \in \mathcal{Y} \implies \sum_{j=1}^{m} d_j y^j$ such that $L(x) = y$, then:
    * $$\begin{aligned} & L(x) = y \\ \implies & L\left(\sum_{i=1}^{n} c_i x^i\right) = \sum_{j=1}^{m} d_j y^j \\ \implies & \sum_{i=1}^{n} c_i L(x^i) = \sum_{j=1}^{m} d_j y^j \\ \implies & \sum_{i=1}^{n} c_i \sum_{j=1}^{m} a_{ji} y^j = \sum_{j=1}^{m} d_j y^j \\ \implies & \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ji} c_i\right) y^j = \sum_{j=1}^{m} d_j y^j \\ \implies & \sum_{i=1}^{n} a_{ji} c_i = d_i \\ \implies & \boldsymbol{A} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} \end{aligned}$$
      - Note the last step uses the uniqueness of coordinate representations

- The key idea is that we can perform a linear transformation between the abstract vector spaces $\mathcal{X}$ and $\mathcal{Y}$ by first going from $\mathcal{X}$ to $\mathbb{R}^n$ using a coordinate representation, then performing the transformation $\mathbb{R}^n \mapsto \mathbb{R}^m$ through a matrix multiplication by $\boldsymbol{A}$ to obtain coordinates for a vector in $\mathcal{Y}$, then mapping back to $\mathcal{Y}$ through the basis
- Note that a transformation has a matrix representation if and only if it is linear and maps between finite dimensional vector spaces

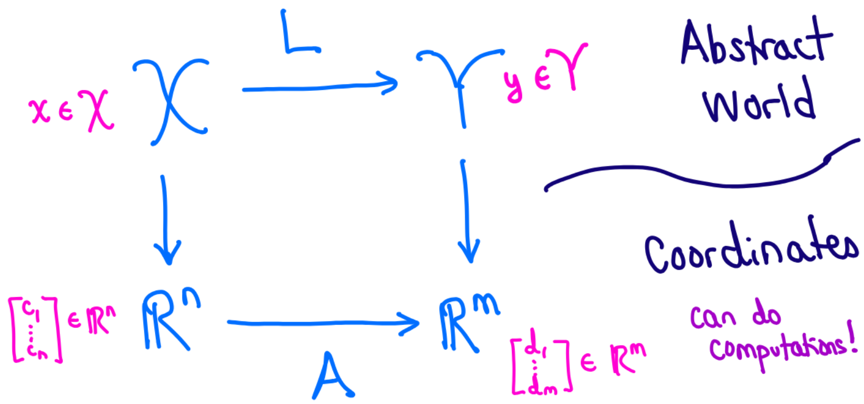

Figure 1: A matrix $\boldsymbol{A}$ as the representation of a linear transformation L between two abstract vector spaces represented with coordinates.

- Example: The matrix representation of a counterclockwise rotation by θ in $\mathbb{R}^2$, using the standard basis, is $\boldsymbol{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$; what is the equivalent transformation, using the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$?
  - Denote the standard basis $\mathcal{E} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$
  - We want to find a matrix $\bar{\boldsymbol{A}}$ that takes us from $\mathbb{R}^2$ represented with $\mathcal{B}$ to another $\mathbb{R}^2$ represented with $\mathcal{B}$; we know that $\boldsymbol{A}$ takes us from $\mathbb{R}^2$ represented with $\mathcal{E}$ to another $\mathbb{R}^2$ represented with $\mathcal{E}$
  - Suppose we can get from basis $\mathcal{B}$ to basis $\mathcal{E}$ through $\boldsymbol{M}$, then we can get back to basis $\mathcal{B}$ by $\boldsymbol{M}^{-1}$
  - Therefore $\bar{\boldsymbol{A}}\boldsymbol{z} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}\boldsymbol{z}$ – first applying $\boldsymbol{M}$ to get to $\mathcal{E}$, then applying $\boldsymbol{A}$ in basis $\mathcal{E}$, and then applying $\boldsymbol{M}^{-1}$ to get back to $\mathcal{B}$
    * Therefore $\bar{\boldsymbol{A}} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$ – a similarity transform
    * Note the order that we write this is kind of reversed
  - Let $\boldsymbol{z}$ have coordinates $\begin{bmatrix} \bar{\xi}_1 \\ \bar{\xi}_2 \end{bmatrix}$ in $\mathcal{B}$ and $\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$ in $\mathcal{E}$, i.e. $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix} \begin{bmatrix} \bar{\xi}_1 \\ \bar{\xi}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$
  - We want to find $\boldsymbol{M}$ such that $\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \boldsymbol{M} \begin{bmatrix} \bar{\xi}_1 \\ \bar{\xi}_2 \end{bmatrix} \implies \boldsymbol{M} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix}$
  - Therefore $\boldsymbol{M} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and we can use this to find $\bar{\boldsymbol{A}}$

Definition

Let $L : \mathcal{X} \mapsto \mathcal{Y}$ be a linear transformation. The *null space* or *kernel* of L is

$$\mathcal{N}(L) = \{\, \boldsymbol{x} \in \mathcal{X} \mid L(\boldsymbol{x}) = \bar{0} \,\}$$

i.e. all the vectors that map to zero. This is a subspace.

The *range* or *image* of L is

$$\mathcal{R}(L) = \{\, \boldsymbol{y} \in \mathcal{Y} \mid \exists \boldsymbol{x} \in \mathcal{X},\, \boldsymbol{y} = L(\boldsymbol{x}) \,\}$$

i.e. all the vectors that can be reached via L. This is another subspace.

- Note for a subspace $\mathcal{V}$ of $\mathcal{X}$, then we denote, in general, the range of $\mathcal{V}$ under a linear transformation L as $L(\mathcal{V}) = \{\, \boldsymbol{y} \in \mathcal{Y} \mid \exists \boldsymbol{x} \in \mathcal{V},\, \boldsymbol{y} = L(x) \,\}$

Definition

Let $L : \mathcal{X} \mapsto \mathcal{Y}$ be a linear transformation between finite dimensional vector spaces $\mathcal{X}, \mathcal{Y}$, then the *rank* of L is defined as

$$\operatorname{rank}(L) = \dim(\mathcal{R}(L))$$

Theorem

$L : \mathcal{X} \mapsto \mathcal{Y}$ for finite dimensional $\mathcal{X}, \mathcal{Y}$ satisfies the following properties:

1. L is injective if and only if $\mathcal{N}(L) = \{\, \bar{0} \,\}$
2. $\dim(\mathcal{R}(L)) + \dim(\mathcal{N}(L)) = \dim(\mathcal{X})$

- The second property (rank-nullity) can be proven as follows:
  - Let $k = \dim(\mathcal{N}(L))$ and $n = \dim(\mathcal{X})$; we want to show $n - k = \dim(\mathcal{R}(L))$
  - Let $\operatorname{span} \boldsymbol{x}^1, \dots, \boldsymbol{x}^k$ be a basis for $\mathcal{N}(L)$, and so $L(\boldsymbol{x}^i) = \bar{0}$ for $i \in [1, k]$
  - Complete the basis such that $\operatorname{span} \boldsymbol{x}^1, \dots, \boldsymbol{x}^k, \boldsymbol{x}^{k+1}, \dots, \boldsymbol{x}^n$ be a basis for $\mathcal{X}$
  - Let $\boldsymbol{x} \in \mathcal{X}$, which has a unique coordinate representation $\boldsymbol{x} = \sum_{i=1}^{n} c_i \boldsymbol{x}^i$ with respect to this basis
  - $$\begin{aligned} L(\boldsymbol{x}) &= L\left(\sum_{i=1}^{n} c_i \boldsymbol{x}^i\right) \\ &= \sum_{i=1}^{k} c_i L(\boldsymbol{x}^i) + \sum_{i=k+1}^{n} c_i L(\boldsymbol{x}^i) \\ &= \sum_{i=k+1}^{n} c_i L(\boldsymbol{x}^i) \end{aligned}$$
  - This suggests $\{\, L(\boldsymbol{x}^i) \,\}$ for $i = k+1, \dots, n$ forms a basis for $\mathcal{R}(L)$
    * To do this, we need to prove that they span $\mathcal{R}(L)$ and that they are linearly independent (in the notes)

Theorem

Let $L : \mathcal{X} \mapsto \mathcal{Y}$, then for any matrix representation $\boldsymbol{A}$ of the linear map L,

$$\dim(\mathcal{R}(L)) = \dim(\mathcal{R}(\boldsymbol{A})) \tag{1}$$
$$\dim(\mathcal{N}(L)) = \dim(\mathcal{N}(\boldsymbol{A})) \tag{2}$$

1. L is surjective if and only if $\text{rank}(\boldsymbol{A}) = \dim(\mathcal{R}(\boldsymbol{A})) = \dim(\mathcal{Y})$, i.e. all rows of $\boldsymbol{A}$ are linearly independent (full row rank)
2. L is injective if and only if $\dim(\mathcal{N}(\boldsymbol{A})) = 0$, i.e. all columns of $\boldsymbol{A}$ are linearly independent (full column rank)
3. L is bijective if and only if $\boldsymbol{A}$ is square and invertible