

# *DSC 140B*

## *Representation Learning*

Lecture 03 | Part 1

**Functions of a Vector**

# Functions of a Vector

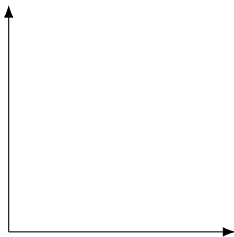
- ▶ In ML, we often work with functions of a vector:  
 $f : \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$ .
- ▶ Example: a prediction function,  $H(\vec{x})$ .
- ▶ Functions of a vector can return:
  - ▶ a number:  $f : \mathbb{R}^d \rightarrow \mathbb{R}^1$
  - ▶ a vector  $\vec{f} : \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$
  - ▶ something else?

# Transformations

- ▶ A **transformation**  $\vec{f}$  is a function that takes in a vector, and returns a vector *of the same dimensionality*.
- ▶ That is,  $\vec{f} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ .

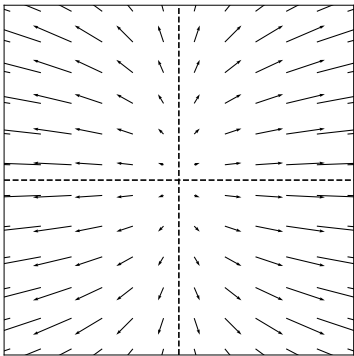
# Visualizing Transformations

- ▶ A transformation is a **vector field**.
  - ▶ Assigns a vector to each point in space.
  - ▶ Example:  $\vec{f}(\vec{x}) = (3x_1, x_2)^T$



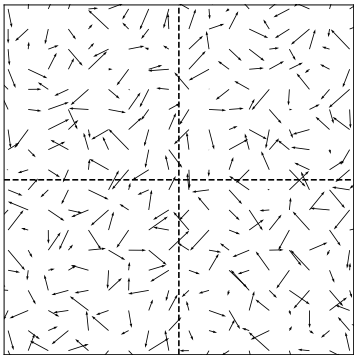
# Example

►  $\vec{f}(\vec{x}) = (3x_1, x_2)^T$



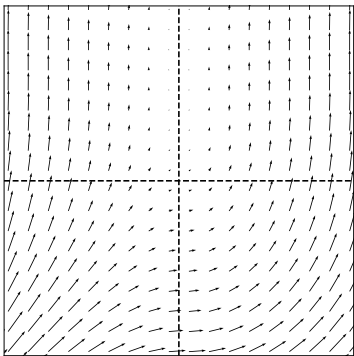
# Arbitrary Transformations

- Arbitrary transformations can be quite complex.



# Arbitrary Transformations

- Arbitrary transformations can be quite complex.



# Linear Transformations

- ▶ Luckily, we often<sup>1</sup> work with simpler, **linear transformations**.
- ▶ A transformation  $f$  is linear if:

$$\vec{f}(\alpha\vec{x} + \beta\vec{y}) = \alpha\vec{f}(\vec{x}) + \beta\vec{f}(\vec{y})$$

---

<sup>1</sup>Sometimes, just to make the math tractable!



# Checking Linearity

- ▶ To check if a transformation is linear, use the definition.
- ▶ **Example:**  $\vec{f}(\vec{x}) = (x_2, -x_1)^T$

## Exercise

Let  $\vec{f}(\vec{x}) = (x_1 + 3, x_2)$ . Is  $\vec{f}$  a linear transformation?

# Implications of Linearity

- Suppose  $\vec{f}$  is a linear transformation. Then:

$$\begin{aligned}\vec{f}(\vec{x}) &= \vec{f}(x_1 \hat{e}^{(1)} + x_2 \hat{e}^{(2)}) \\ &= x_1 \vec{f}(\hat{e}^{(1)}) + x_2 \vec{f}(\hat{e}^{(2)})\end{aligned}$$

- I.e.,  $\vec{f}$  is **totally determined** by what it does to the basis vectors.

# The **Complexity** of Arbitrary Transformations

- ▶ Suppose  $f$  is an **arbitrary** transformation.
- ▶ I tell you  $\vec{f}(\hat{e}^{(1)}) = (2, 1)^T$  and  $\vec{f}(\hat{e}^{(2)}) = (-3, 0)^T$ .
- ▶ I tell you  $\vec{x} = (x_1, x_2)^T$ .
- ▶ What is  $\vec{f}(\vec{x})$ ?

# The **Simplicity** of Linear Transformations


- ▶ Suppose  $f$  is a **linear** transformation.
- ▶ I tell you  $\vec{f}(\hat{e}^{(1)}) = (2, 1)^T$  and  $\vec{f}(\hat{e}^{(2)}) = (-3, 0)^T$ .
- ▶ I tell you  $\vec{x} = (x_1, x_2)^T$ .
- ▶ What is  $\vec{f}(\vec{x})$ ?

## Exercise

- ▶ Suppose  $f$  is a **linear** transformation.
- ▶ I tell you  $\vec{f}(\hat{e}^{(1)}) = (2, 1)^T$  and  $\vec{f}(\hat{e}^{(2)}) = (-3, 0)^T$ .
- ▶ I tell you  $\vec{x} = (3, -4)^T$ .
- ▶ What is  $\vec{f}(\vec{x})$ ?

## Key Fact

- ▶ Linear functions are determined **entirely** by what they do on the basis vectors.
- ▶ I.e., to tell you what  $f$  does, I only need to tell you  $\vec{f}(\hat{e}^{(1)})$  and  $\vec{f}(\hat{e}^{(2)})$ .
- ▶ This makes the math easy!

A photograph of a formal garden, likely the gardens of Stourhead in England. The garden features a central lawn area, symmetrical hedges, and various flower beds. The background is a dense forest of tall trees. The text "Arbitrary Transformations" is overlaid in the upper right quadrant.

Arbitrary  
Transformations

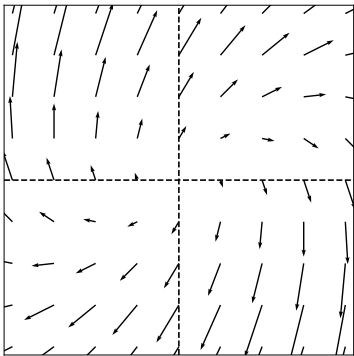
Linear  
Transformations

A small, stylized logo consisting of a lowercase letter 'b' inside a square frame, located in the bottom left corner of the image.



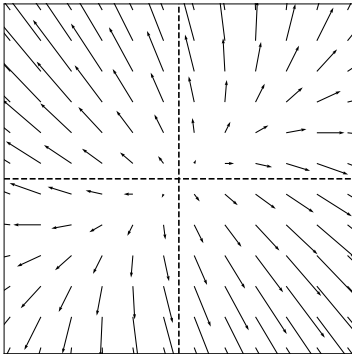
# Example Linear Transformation

►  $\vec{f}(\vec{x}) = (x_1 + 3x_2, -3x_1 + 5x_2)^T$



# Another Example Linear Transformation

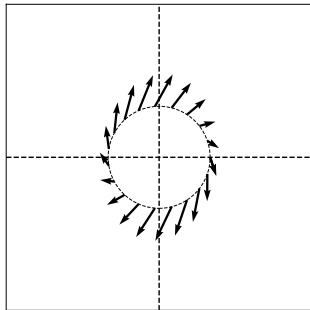
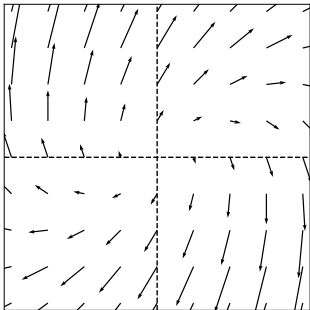
►  $\vec{f}(\vec{x}) = (2x_1 - x_2, -x_1 + 3x_2)^T$

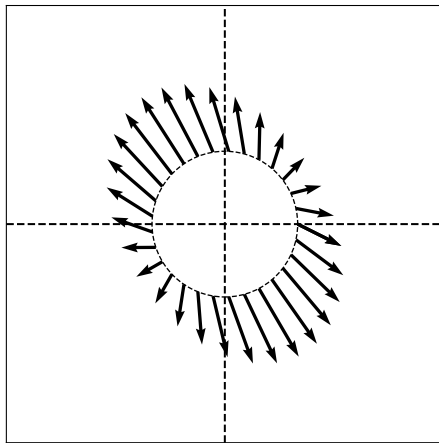
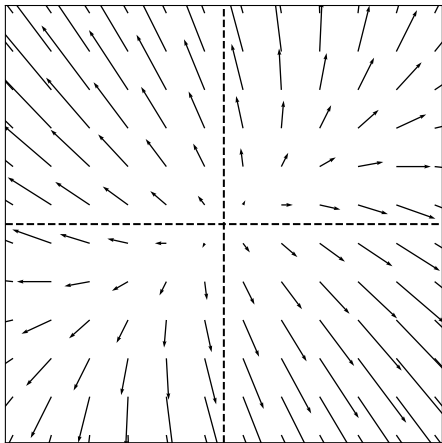


# Note

- Because of linearity, along any given direction  $\vec{f}$  changes only in scale.

$$\vec{f}(\lambda \hat{x}) = \lambda \vec{f}(\hat{x})$$





# Linear Transformations and Bases

- ▶ We have been writing transformations in coordinate form. For example:

$$\vec{f}(\vec{x}) = (x_1 + x_2, x_1 - x_2)^T$$

- ▶ To do so, we assumed the **standard basis**.
- ▶ If we use a different basis, the formula for  $\vec{f}$  changes.

# Example

- ▶ Suppose that in the standard basis,  $\vec{f}(\vec{x}) = (x_1 + x_2, x_1 - x_2)^T$ .
- ▶ Let  $\hat{u}^{(1)} = \frac{1}{\sqrt{2}}(1, 1)^T$  and  $\hat{u}^{(2)} = \frac{1}{\sqrt{2}}(-1, 1)^T$ .
- ▶ Write  $[\vec{x}]_{\mathcal{U}} = (z_1, z_2)^T$ .
- ▶ What is  $[\vec{f}(\vec{x})]_{\mathcal{U}}$  in terms of  $z_1$  and  $z_2$ ?

# DSC 140B

## Representation Learning

Lecture 03 | Part 2

**Matrices**

# Matrices?

- ▶ I thought this week was supposed to be about linear algebra... Where are the matrices?



# Matrices?

- ▶ I thought this week was supposed to be about linear algebra... Where are the matrices?
- ▶ What is a matrix, anyways?

# What is a matrix?

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

# Recall: Linear Transformations

- ▶ A **transformation**  $\vec{f}(\vec{x})$  is a function which takes a vector as input and returns a vector of the same dimensionality.
- ▶ A transformation  $\vec{f}$  is **linear** if

$$\vec{f}(\alpha \vec{u} + \beta \vec{v}) = \alpha \vec{f}(\vec{u}) + \beta \vec{f}(\vec{v})$$

# Recall: Linear Transformations

- ▶ **Key** consequence of **linearity**: to compute  $\vec{f}(\vec{x})$ , only need to know what  $\vec{f}$  does to basis vectors.
- ▶ Example:

$$\vec{x} = 3\hat{e}^{(1)} - 4\hat{e}^{(2)} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$$

$$\vec{f}(\hat{e}^{(1)}) = -\hat{e}^{(1)} + 3\hat{e}^{(2)}$$

$$\vec{f}(\hat{e}^{(2)}) = 2\hat{e}^{(1)}$$

$$\vec{f}(\vec{x}) =$$

# Matrices

- ▶ **Idea:** Since  $\vec{f}$  is defined by what it does to basis, place  $\vec{f}(\hat{e}^{(1)})$ ,  $\vec{f}(\hat{e}^{(2)})$ , ... into a table as columns
- ▶ This is the **matrix** representing<sup>2</sup>  $\vec{f}$

$$\begin{aligned}\vec{f}(\hat{e}^{(1)}) &= -\hat{e}^{(1)} + 3\hat{e}^{(2)} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \\ \vec{f}(\hat{e}^{(2)}) &= 2\hat{e}^{(1)} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}\end{aligned}\qquad \begin{pmatrix} -1 & 2 \\ 3 & 0 \end{pmatrix}$$

---

<sup>2</sup>with respect to the standard basis  $\hat{e}^{(1)}, \hat{e}^{(2)}$

## Exercise

Write the matrix representing  $\vec{f}$  with respect to the standard basis, given:

$$\vec{f}(\hat{e}^{(1)}) = (1, 4, 7)^T$$

$$\vec{f}(\hat{e}^{(2)}) = (2, 5, 7)^T$$

$$\vec{f}(\hat{e}^{(3)}) = (3, 6, 9)^T$$

## Exercise

Suppose  $\vec{f}$  has the matrix below:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Let  $\vec{x} = (-2, 1, 3)^T$ . What is  $\vec{f}(\vec{x})$ ?

## Main Idea

A square ( $n \times n$ ) matrix can be interpreted as a compact representation of a linear transformation  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ .



# What is matrix multiplication?

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} \phantom{0} \\ \phantom{0} \\ \phantom{0} \end{pmatrix}$$

## A low-level definition

$$(A\vec{x})_i = \sum_{j=1}^n A_{ij}x_j$$

## A low-level interpretation

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} = -2 \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} + 1 \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} + 3 \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$$

**In general...**

$$\begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \vec{a}^{(1)} & \vec{a}^{(2)} & \vec{a}^{(3)} \\ \downarrow & \downarrow & \downarrow \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \vec{a}^{(1)} + x_2 \vec{a}^{(2)} + x_3 \vec{a}^{(3)}$$

# Matrix Multiplication

$$\vec{x} = x_1 \hat{e}^{(1)} + x_2 \hat{e}^{(2)} + x_3 \hat{e}^{(3)} = (x_1, x_2, x_3)^T$$
$$\vec{f}(\vec{x}) = x_1 \vec{f}(\hat{e}^{(1)}) + x_2 \vec{f}(\hat{e}^{(2)}) + x_3 \vec{f}(\hat{e}^{(3)})$$

$$A = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \vec{f}(\hat{e}^{(1)}) & \vec{f}(\hat{e}^{(2)}) & \vec{f}(\hat{e}^{(3)}) \\ \downarrow & \downarrow & \downarrow \end{pmatrix}$$
$$A\vec{x} = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \vec{f}(\hat{e}^{(1)}) & \vec{f}(\hat{e}^{(2)}) & \vec{f}(\hat{e}^{(3)}) \\ \downarrow & \downarrow & \downarrow \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$
$$= x_1 \vec{f}(\hat{e}^{(1)}) + x_2 \vec{f}(\hat{e}^{(2)}) + x_3 \vec{f}(\hat{e}^{(3)})$$

# Matrix Multiplication

- ▶ Matrix  $A$  represents a linear transformation  $\vec{f}$ 
  - ▶ With respect to the standard basis
  - ▶ If we use a different basis, the matrix changes!
- ▶ Matrix multiplication  $A\vec{x}$  **evaluates**  $\vec{f}(\vec{x})$

## **What are they, *really*?**

- ▶ Matrices are sometimes just tables of numbers.
- ▶ But they often have a deeper meaning.

## Main Idea

A square ( $n \times n$ ) matrix can be interpreted as a compact representation of a linear transformation  $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ .

What's more, if  $A$  represents  $\vec{f}$ , then  $A\vec{x} = \vec{f}(\vec{x})$ ; that is, multiplying by  $A$  is the same as evaluating  $\vec{f}$ .



# Example

$$\vec{x} = 3\hat{e}^{(1)} - 4\hat{e}^{(2)} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$$

$$A =$$

$$\vec{f}(\hat{e}^{(1)}) = -\hat{e}^{(1)} + 3\hat{e}^{(2)}$$

$$\vec{f}(\hat{e}^{(2)}) = 2\hat{e}^{(1)}$$

$$\vec{f}(\vec{x}) =$$

$$A\vec{x} =$$

## Note

- ▶ All of this works because we assumed  $\vec{f}$  is **linear**.
- ▶ If it isn't, evaluating  $\vec{f}$  isn't so simple.

## Note

- ▶ All of this works because we assumed  $\vec{f}$  is **linear**.
- ▶ If it isn't, evaluating  $\vec{f}$  isn't so simple.
- ▶ Linear algebra = simple!

# Matrices in Other Bases

- The matrix of a linear transformation wrt the **standard basis**:

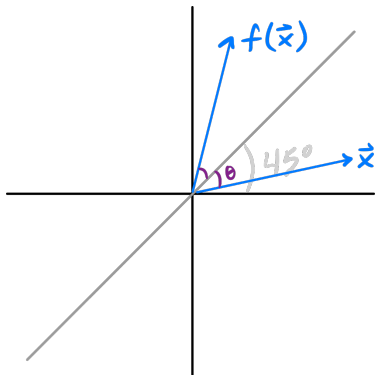
$$\begin{pmatrix} \uparrow & \uparrow & \uparrow & \\ \vec{f}(\hat{e}^{(1)}) & \vec{f}(\hat{e}^{(2)}) & \cdots & \vec{f}(\hat{e}^{(d)}) \\ \downarrow & \downarrow & \downarrow & \end{pmatrix}$$

- With respect to basis  $\mathcal{U}$ :

$$\begin{pmatrix} \uparrow & \uparrow & \uparrow & \\ [\vec{f}(\hat{u}^{(1)})]_{\mathcal{U}} & [\vec{f}(\hat{u}^{(2)})]_{\mathcal{U}} & \cdots & [\vec{f}(\hat{u}^{(d)})]_{\mathcal{U}} \\ \downarrow & \downarrow & \downarrow & \end{pmatrix}$$

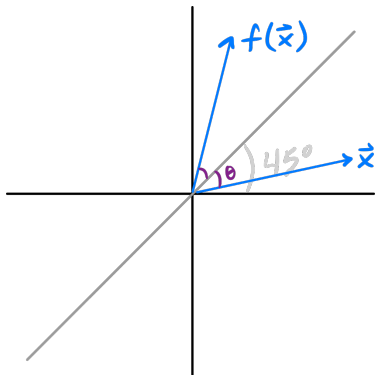
# Matrices in Other Bases

- Consider the transformation  $\vec{f}$  which “mirrors” a vector over the line of  $45^\circ$ .



- What is its matrix in the standard basis?

# Matrices in Other Bases



- ▶ Let  $\hat{u}^{(1)} = \frac{1}{\sqrt{2}}(1, 1)^T$
- ▶ Let  $\hat{u}^{(2)} = \frac{1}{\sqrt{2}}(-1, 1)^T$
- ▶ What is  $[\vec{f}(\hat{u}^{(1)})]_{\mathcal{U}}$ ?
- ▶  $[\vec{f}(\hat{u}^{(2)})]_{\mathcal{U}}$ ?
- ▶ What is the matrix?