# Data 345

Applied Linear Algebra for Statistical Learning

Class 9 (Sep. 25, 2025)

# Functions / Mappings

- Given sets $X$ and $Y$, a function or mapping $f$ from $X$ to $Y$ is a rule that assigns each element of $X$ to precisely one element in $Y$.

Set of inputs (*domain*)

$$f : X \to Y$$

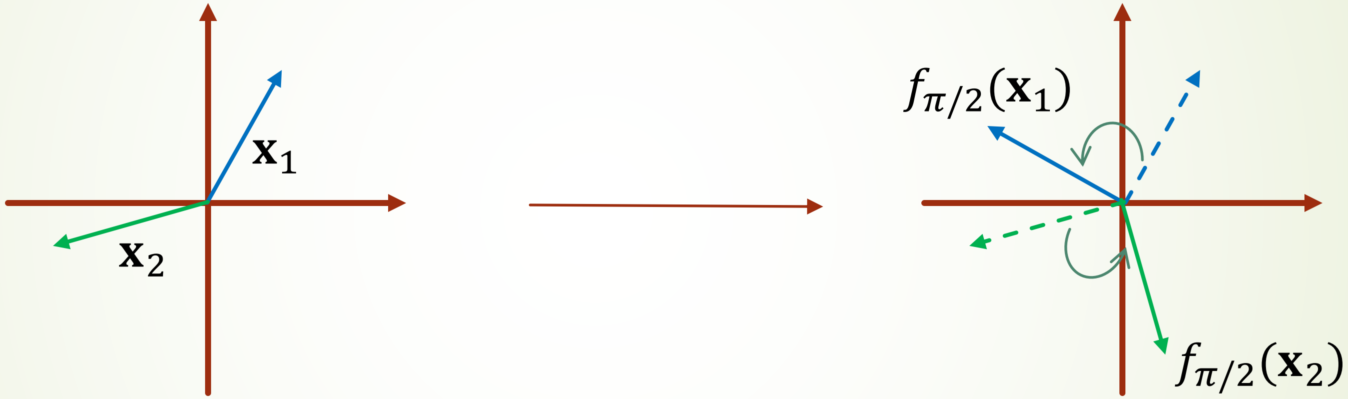
Name of function

Set of possible outputs (*codomain*)

- Formally, a function requires no structure other than two sets. However, we are often interested in special kinds of functions which preserve a particular structure or have nice properties.
- Example: A *continuous* function from calculus preserves local distortion of distance (i.e., a small perturbation of an input results in a "small" perturbation of the output)

# Functions / Mappings

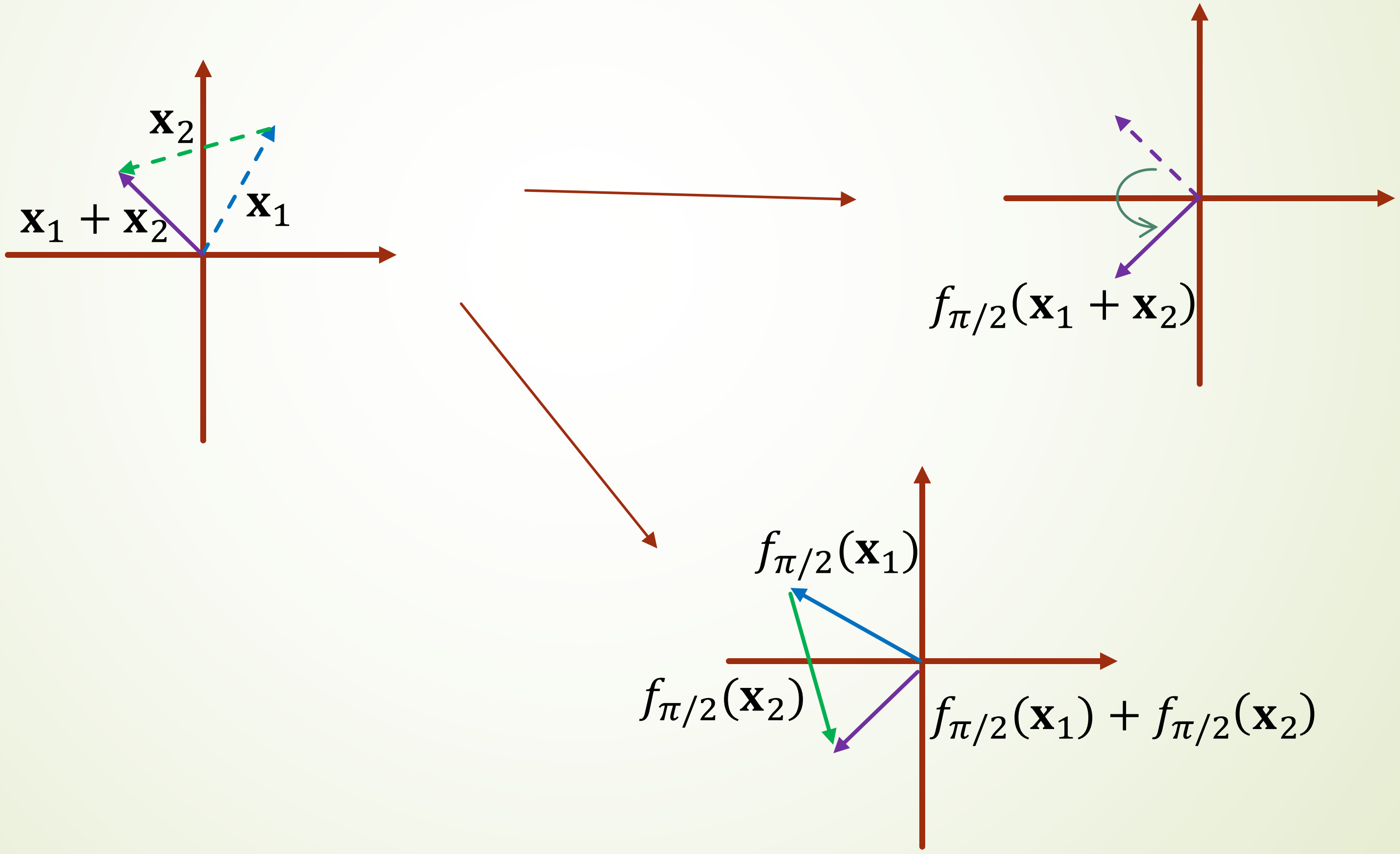
- Let's consider an example that is common in linear algebra. Let $f_{\pi/2}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a function that rotates every vector by $\frac{\pi}{2}$ radians (counterclockwise) about the origin.




- This is a relatively simple but important transformation. How does this function "interact" with the vector space structure on $\mathbb{R}^2$?

# Functions / Mappings

- If we have two vectors $x_1$ and $x_2$ then we can either add before or after rotating and get the same result.



# Functions / Mappings

- Similarly, since scalar multiplication just lengthens/shortens a vector in the same direction, we may perform scalar multiplication either before or after rotating and still get the same result.
- This brings out two very important properties of $f_{\pi/2}$:
  - $f_{\pi/2}$ preserves the operation of vector addition: for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$, we have that $f_{\pi/2}(\mathbf{x}_1 + \mathbf{x}_2) = f_{\pi/2}(\mathbf{x}_1) + f_{\pi/2}(\mathbf{x}_2)$.
  - $f_{\pi/2}$ preserves the operation of scalar multiplication: for any $x \in \mathbb{R}^2$ and any $\lambda \in \mathbb{R}$, we have that $f_{\pi/2}(\lambda\mathbf{x}) = \lambda \cdot f_{\pi/2}(\mathbf{x})$.
- In other words, the function $f_{\pi/2}$ does not break any of the vector space structure to which we have grown accustomed.
- It is hopefully easy to see that the angle $\pi/2$ is not important here – a function rotating by any angle $\theta$ would satisfy the same properties.

# Linear Transformations

- Let $V$ and $W$ be any two real vector spaces. A mapping $f: V \to W$ is said to be a **linear transformation** or **linear mapping** if:
  - For any two $\mathbf{x}, \mathbf{y} \in V$, $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$.
  - For any real $\lambda$ and any $\mathbf{x} \in V$, $f(\lambda\mathbf{x}) = \lambda f(\mathbf{x})$.
- Example: Define $i = \sqrt{-1}$. Then a **complex number** is any expression of the form $a + bi$, where $a$ and $b$ are real numbers. The set of all complex numbers is denoted by $\mathbb{C}$.
- $\mathbb{C}$ has a (real) vector space structure: $(a + bi) + (c + di) = (a + c) + (b + d)i$ (adding and collecting like terms), and $\lambda(a + bi) = \lambda a + \lambda bi$.
- Define a mapping $f: \mathbb{R}^2 \to \mathbb{C}$ where $f([x_1, x_2]) = x_1 + x_2 i$. Then, $f$ is a linear transformation.

# Linear Transformations & Matrices

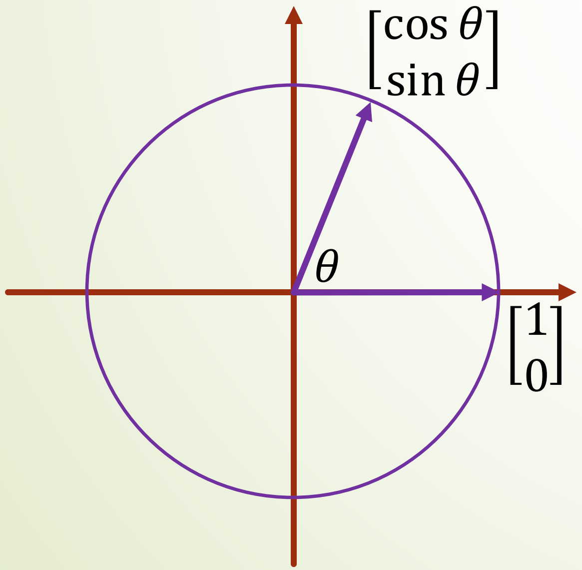
- Let's return to the transformation $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ (this time, we will rotate by an arbitrary angle $\theta$).
- We know the effect of this transformation, but it's not super clear how we could write down a "formula" for it.
- In other words, if we have $\mathbf{x} = [x_1, x_2]^T$, we would like to be able to write down $f_\theta(\mathbf{x}) = [y_1, y_2]^T$ where $y_1$ and $y_2$ are both expressions involving $x_1$ and $x_2$.
- Remember that in $\mathbb{R}^2$ we have a *standard basis* $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$ which allows us to write any vector in $\mathbb{R}^2$ as a linear combination of these basis vectors.
- If $\mathbf{x} = \begin{bmatrix}x_1\\x_2\end{bmatrix}$ then $\mathbf{x} = x_1 \cdot \begin{bmatrix}1\\0\end{bmatrix} + x_2 \cdot \begin{bmatrix}0\\1\end{bmatrix}$.

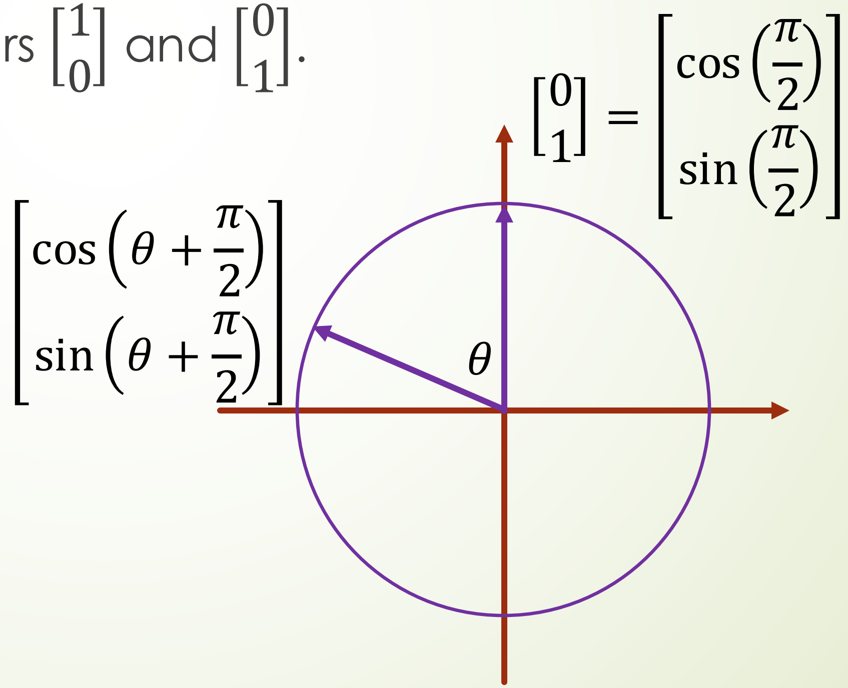
# Linear Transformations & Matrices

- Knowing (or being willing to believe) that $f_\theta$ is linear allows us to say:

$$f_\theta(\mathbf{x}) = f_\theta\left(x_1\begin{bmatrix}1\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\end{bmatrix}\right) = f_\theta\left(x_1\begin{bmatrix}1\\0\end{bmatrix}\right) + f_\theta\left(x_2\begin{bmatrix}0\\1\end{bmatrix}\right) = x_1 f_\theta\left(\begin{bmatrix}1\\0\end{bmatrix}\right) + x_2 f_\theta\left(\begin{bmatrix}0\\1\end{bmatrix}\right).$$

- So, to determine a "formula" for $f_\theta$ we really only need to establish what $f_\theta$ does to the basis vectors $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$.



# Linear Transformations & Matrices

- The quantities $\cos\left(\theta + \frac{\pi}{2}\right)$ and $\sin\left(\theta + \frac{\pi}{2}\right)$ are very ugly, so let's simplify…
- $\cos\left(\theta + \frac{\pi}{2}\right) = \cos\theta\cos\pi/2 - \sin\theta\sin\pi/2$, so $\cos\left(\theta + \frac{\pi}{2}\right) = -\sin\theta$.
- $\sin\left(\theta + \frac{\pi}{2}\right) = \sin\theta\cos\pi/2 + \cos\theta\sin\pi/2$, so $\sin\left(\theta + \frac{\pi}{2}\right) = \cos\theta$.
- Thus, $f_\theta\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}$ and $f_\theta\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-\sin\theta\\\cos\theta\end{bmatrix}$.
- This means we may more generally compute:

$$
\begin{aligned}
f_\theta(\mathbf{x}) &= x_1 f_\theta\left(\begin{bmatrix}1\\0\end{bmatrix}\right) + x_2 f_\theta\left(\begin{bmatrix}0\\1\end{bmatrix}\right) \\
&= x_1\begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix} + x_2\begin{bmatrix}-\sin\theta\\\cos\theta\end{bmatrix} \\
&= \begin{bmatrix}x_1\cos\theta - x_2\sin\theta\\x_1\sin\theta + x_2\cos\theta\end{bmatrix}
\end{aligned}
$$

# Linear Transformations & Matrices

- Wait! If you've been half paying attention, you might notice:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1\cos\theta - x_2\sin\theta \\ x_1\sin\theta + x_2\cos\theta \end{bmatrix}$$

- This means we encoded the function $f_\theta$ as a **matrix**!
- Now comes the punchline: **every linear transformation may be encoded as a matrix** in a similar way.
- Specifically, let $T: V \to W$ be a linear transformation from $V$ to $W$, where $\dim(V) = n$ and $\dim(W) = m$. Let $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ and $C = (\mathbf{c}_1, \dots, \mathbf{c}_m)$ be ordered bases for $V$ and $W$, respectively.
- For any $j \in \{1, \dots, n\}$, $T(\mathbf{b_j})$ can be represented as a linear combination of $\mathbf{c}_1$ through $\mathbf{c}_m$, i.e.,

$$T(\mathbf{b}_j) = \alpha_{1j}\mathbf{c}_1 + \cdots + \alpha_{mj}\mathbf{c}_m \longrightarrow A_T = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}$$

# Linear Transformations & Matrices

- The matrix $A_T$ is called the **transformation matrix** of $T$ with respect to $B$ and $C$.
- Since linear transformations are entirely determined by their effect on basis elements, a transformation essentially just summarizes the effect of $T$ on each individual basis element from $B$ in terms of the basis $C$.
- Example: Define a linear mapping $T: \mathbb{R}^3 \to \mathbb{R}^4$ as:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 + 2x_2 + x_3 \\ x_1 + x_2 + x_3 \\ x_1 - 3x_2 \\ 2x_1 + 3x_2 + x_3 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 1 \\ 1 \\ 2 \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \\ -3 \\ 3 \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \qquad A_T = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & -3 & 0 \\ 2 & 3 & 1 \end{bmatrix}$$

# Image & Kernel

- Let $T: V \to W$ be a linear transformation. The **image** or **range** of $T$ is simply $T(V)$. The **kernel** or **null space** of $T$ is the set of all vectors $v$ in $V$ such that $T(v) = \mathbf{0}_W$ -- in other words, the kernel of $T$ is the set of vectors in $V$ that get annihilated (sent to the zero vector) by the transformation $T$.
- Because $T$ is linear, the image of $T$ is always a subspace of $W$, while the kernel of $T$ is always a subspace of $V$.