Data 345

Applied Linear Algebra for Statistical Learning

Class 11 (Oct. 2, 2025)

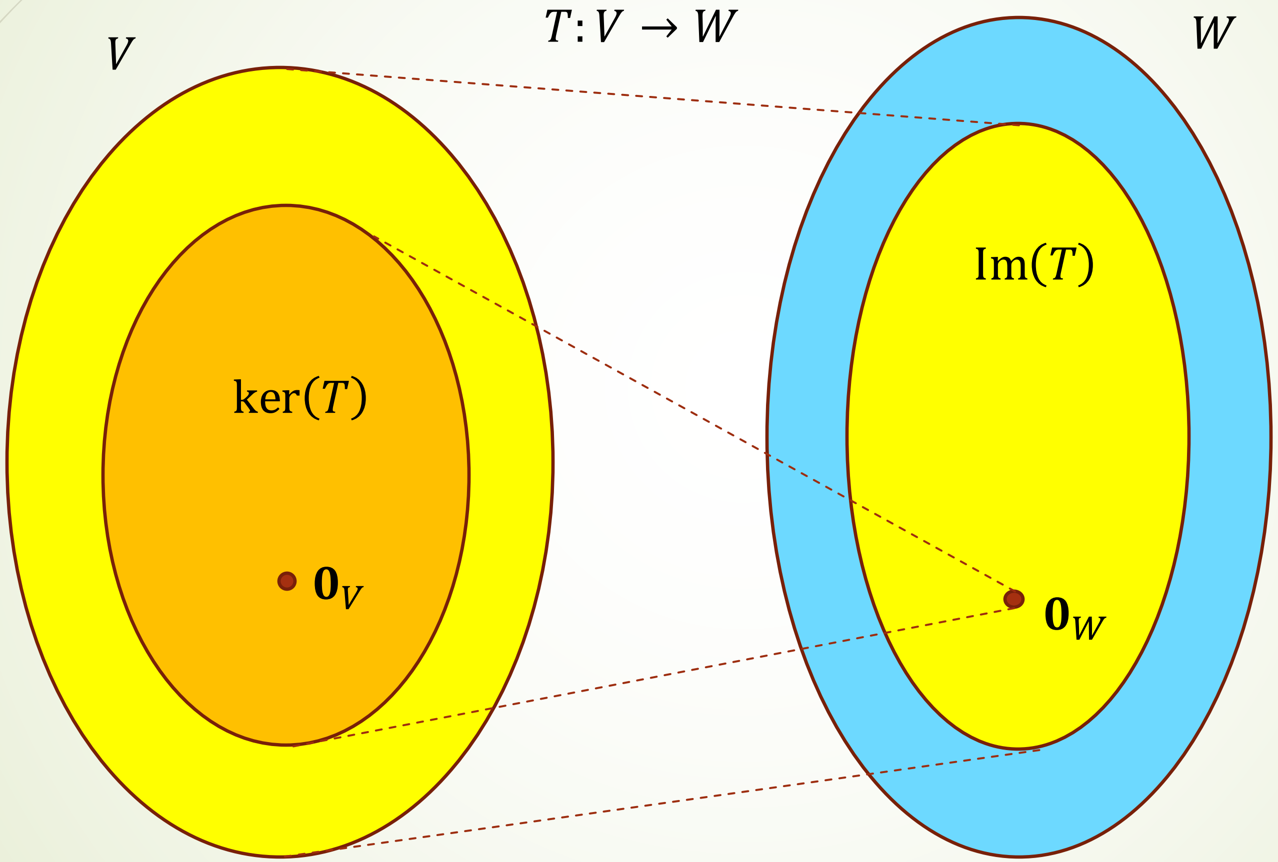
Reminder

- Today's my last class with y'all! As of next week, you will be picking up the class with a new instructor and new modality.
- **Old:** In-person M/Th 12:00 – 1:50
- **New:** Online synchronous Th 12:00 – 1:50 / Online asynchronous otherwise

Image & Kernel

- Let $T: V \to W$ be a linear transformation. The **image** or **range** of T is simply $T(V)$. The **kernel** of T is the set of all vectors $\mathbf{v}$ in V such that $T(\mathbf{v}) = \mathbf{0}_W$.
- 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. The image of T is often written $\mathrm{Im}(T)$, while the kernel is often written $\ker(T)$.
- For any linear transformation $T: V \to W$, we have that $T(\mathbf{0}_V) = \mathbf{0}_W$ so the null space / kernel is never empty.
- 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.

Image & Kernel



The kernel $\ker(T)$ is the part of the space V that is squashed into the zero vector.

The image $\mathrm{Im}(T)$ is the version of V that has been successfully "transferred" to W.

Image & Kernel

- If $T\colon V \to W$ is a linear transformation and $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ is an $m \times n$ transformation matrix for T (where $\mathbf{a}_1$ through $\mathbf{a}_n$ are its columns) then $\mathrm{Im}(T) = \mathrm{span}[\mathbf{a}_1, \dots, \mathbf{a}_n]$. In particular, $\dim\big(\mathrm{Im}(T)\big) = \mathrm{rank}(A)$. This is why we also sometimes call $\mathrm{Im}(T)$ the **column space** of A, written $\mathrm{col(A)}$.
- Similarly, the **null space** of a matrix A, written $\mathrm{null}(A)$, is the subspace of V whose basis vectors are the solution to the equation $A\mathbf{x} = \mathbf{0}_W$. The kernel of a linear transformation $\ker(T)$ is the same as $\mathrm{null}(A)$.
- Oftentimes the convention is to use $\mathrm{Im}(T)$ and $\ker(T)$ when referring to the transformation T (which is globally defined on V without considering basis) but use $\mathrm{col}(A)$ and $\mathrm{null}(A)$ when referring to the matrix A (which depends on a choice of ordered basis for V and W).
- The space $\mathrm{null}(A)$ focuses on the relationship among columns – in particular, whether the columns are linearly independent.

Image & Kernel

- Example: Let $T: \mathbb{R}^4 \to \mathbb{R}^2$ be the linear mapping defined as

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix}.$$

- Taking T with respect to the standard basis on $\mathbb{R}^4$ and $\mathbb{R}^2$ gives:

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad T\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

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

Therefore: $\mathrm{Im}(T) = \mathrm{span}\left[\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right]$

Image & Kernel

- To calculate $\mathrm{null}(A)$, we row reduce:

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

- Then since column 3 is (-1/2) times column 2, we get $\mathbf{0} = \frac{1}{2}\mathbf{a}_2 + \mathbf{a}_3$ and similarly $\mathbf{0} = \mathbf{a}_1 - \frac{1}{2}\mathbf{a}_2 - \mathbf{a}_4$. This means that

$$\mathrm{null}(A) = \mathrm{span}\left[\begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}\right]$$

Rank-Nullity Theorem

- Perhaps you noticed that there's some relationship between matrix rank (the number of columns with pivots) and the number of free variables (which determines the basis for the null space).
- This relationship is characterized by the "Rank-Nullity Theorem," which says that if $T: V \to W$ is a linear transformation then

$$\dim(\ker(T)) + \dim(\mathrm{Im}(T)) = \dim V.$$

- When $\dim(V) < \infty$, representing T with a matrix A gives

$$\mathrm{null}(A) + \mathrm{rank}(A) = \dim(V).$$

Rank-Nullity Theorem

- Generally, a function $f: X \to Y$ is said to be **surjective** or **onto** if for every y in Y there exists some $x_y \in X$ for which $f(x_y) = y$. For linear algebra, a linear transformation $T: V \to W$ is surjective if and only if $\dim\big(\mathrm{Im}(T)\big) = \dim(W)$.
- This means that a matrix $A \in \mathbb{R}^{m \times n}$ represents a surjective linear transformation if and only if $\mathrm{rank}(A) = m$.
- A related notion is that a function $f: X \to Y$ is called **injective** or **one-to-one** if whenever $f(x_1) = f(x_2)$ we have that $x_1 = x_2$ (in other words f never repeats the same output). In linear algebra a linear transformation $T: V \to W$ is one-to-one if and only if $\ker(T) = \{\mathbf{0}\}$.
- A **linear isomorphism** $T: V \to W$ is a linear transformation that is both one-to-one and onto.
- Thanks to the Rank-Nullity Theorem we get the amazing fact that a linear transformation $T: V \to W$ (where $\dim V = \dim W$) is surjective if and only if it is injective, meaning the only onto maps between same-dimension spaces are one-to-one.