Lecture 15: Similarity

March 5, 2024

The Rank-Nullity Theorem

The Rank-Nullity Theorem relates the dimensions of the kernel, image and domain of a linear transformation. The dimension of the image of a linear trasformation is called its *rank*, and the dimension of the kernel is called the *nullity*. The rank of T is equal to the rank of the matrix of T, since the image of T is the column space of this matrix.

Theorem (Rank-Nullity Theorem) Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, where V and W are finite-dimensional vector spaces over a field $\mathbb{F}$. Then

$$\dim(\ker T) + \operatorname{rank} T = n.$$

Informally, the Rank-Nullity Theorem says that the full dimension of the domain must be accounted for in the combination of the kernel and the image.

Proof of the Rank-Nullity Theorem

$T : \mathbb{R}^n \to \mathbb{R}^m$ a linear transformation. $\boxed{\dim(\ker T) + \operatorname{rank} T = n.}$

1. Write k for $\dim(\ker T)$ and let $\{b_1, \dots, b_k\}$ be a basis of $\ker T$.
2. Extend this to a basis $\{b_1, \dots, b_k, c_{k+1} \dots, c_n\}$ of $\mathbb{R}^n$.
3. Since T sends each b_i to 0, the image under T of every element of $\mathbb{R}^n$ is a linear combination of $T(c_{k+1}), \dots, T(c_n)$.
4. Also $\{T(c_{k+1}), \dots, T(c_n)\}$ is a linearly independent subset of $\mathbb{R}^m$. To see this, suppose for some scalars $a_{k+1}, \dots, a_n$ that $a_{k+1}T(c_{k+1}) + a_{k+2}T(c_{k+2}) + \cdots + a_nT(c_n) = 0$. Then

   $$a_{k+1}c_{k+1} + \cdots + a_nc_n \in \ker T \implies a_{k+1}c_{k+1} + \cdots + a_nc_n \in \langle b_1, \dots, b_k \rangle.$$

   Since $\{b_1, \dots, b_k, c_{k+1} \dots, c_n\}$ is linearly independent in $\mathbb{R}^n$, this means that $a_{k+1}c_{k+1} + a_{k+2}c_{k+2} + \cdots + a_nc_n = 0$, and each $a_j = 0$.
5. It follows that $\{T(c_{k+1}), \dots, T(c_n)\}$ is a basis for the image of T, so this image has dimension $n - k$, as required.

Linear transformations and change of basis

Definition Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, and let $B = \{b_1, \dots, b_n\}$ be a basis of $\mathbb{R}^n$. The matrix of T with respect to B is the $n \times n$ matrix that has the B-coordinates of $T(b_1), T(b_2), \dots, T(b_n)$ as its n columns. This matrix M satisfies

$$\boxed{[T(v)]_B = M[v]_B, \text{ for all } v \in \mathbb{R}^n}$$

Example Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by $v \to Av$, where

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

Let B be the (ordered) basis of $\mathbb{R}^3$ with elements

$$b_1 = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}, \ b_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \ b_3 = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}.$$

What is the matrix A' of T with respect to B?

A diagonal representation

$$T(b_1) = \begin{bmatrix} -2 & 2 & 1 \\ 4 & 5 & -1 \\ -4 & -8 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix} = 2b_1 \implies [T(b_1)]_B = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

$$T(b_2) = \begin{bmatrix} -2 & 2 & 1 \\ 4 & 5 & -1 \\ -4 & -8 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -6 \\ 3 \\ 0 \end{bmatrix} = -3b_2 \implies [T(b_2)]_B = \begin{bmatrix} 0 \\ -3 \\ 0 \end{bmatrix}$$

$$T(b_3) = \begin{bmatrix} -2 & 2 & 1 \\ 4 & 5 & -1 \\ -4 & -8 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \\ 14 \end{bmatrix} = 7b_3 \implies [T(b_3)]_B = \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix}$$

The matrix A' of T with respect to B is diagonal. For describing this transformation T, B is a better basis than the standard one.

$$A' = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

This means: for any $v \in \mathbb{R}^3$,

$$\underbrace{[T(v)]_B}_{B-\text{coordinates of } T(v)} = \underbrace{A'[v]_B}_{\text{matrix}-\text{vector product}} .$$

Change of basis again

Let P be the matrix with the basis vectors from B as columns.

From Lecture 14, P^{-1} is the change of basis matrix from the standard basis to B. For any element v of $\mathbb{R}^3$, its B-coordinates are given by the matrix-vector product

$$[v]_B = P^{-1}v.$$

Equivalently, if we start with the B-coordinates, then the standard coordinates of v are given by

$$v = P[v]_B.$$

So P itself is the change of basis matrix from B to the standard basis.

Similarity - the relation of A and A'

Starting with A, the matrix of $T : \mathbb{R}^3 \to \mathbb{R}^3$ with respect to the standard basis, how do we find A' the matrix of T with respect to B?

1. Take a vector v of $\mathbb{R}^3$, written in B-coordinates as the column $[v]_B$.
2. Convert to standard coordinates (so that we can apply T by multiplying by A): take the product $P[v]_B$
3. Apply T: left-multiply by A to get $AP[v]_B$. This column has the standard coordinates of $T(v)$.
4. Convert to B-coordinates: left-multiply by P^{-1}, the change of basis matrix from standard to B. This gives $P^{-1}AP[v]_B$. This column has the B-coordinates of $T(v)$.
5. Conclusion: For any element v of $\mathbb{R}^3$, the B-coordinates of $T(v)$ are given by $(P^{-1}AP)[v]_B$.

The B-matrix of T is $P^{-1}AP$, where P has the elements of B as columns.

Similar Matrices

Definition Two square matrices A and B are similar if $B = P^{-1}AP$ for an invertible matrix P.

Notes

1. Two distinct matrices are similar if and only if they represent the same linear transformation, with respect to different bases.
2. We can't tell by glancing at a pair of square matrices if they are similar or not, but there is one feature that is easy to check. The trace of a square matrix is the sum of the entries on the main diagonal, from top left to bottom right. If two matrices are similar, they have the same trace.
3. Similar matrices also have some other features in common, such as having the same determinant (more on that later).
4. Our example showed that the 3×3 matrix $A = \begin{bmatrix} -2 & 2 & 1 \\ 4 & 5 & -1 \\ -4 & -8 & 3 \end{bmatrix}$ is similar to the diagonal matrix $\mathrm{diag}(2, -3, 7)$. We say A is diagonalizable in this situation.

Two equivalent interpretations of diagonalizability

1. From the diagonal form of A' we have $T(b_1) = 2b_1$, $T(b_2) = -3b_2$ and $T(b_3) = 7b_3$. This means that each of the basis elements b_1, b_2, b_3 is mapped by T to a scalar multiple of itself - each of them is an *eigenvector* of T.
2. We can rearrange the version $P^{-1}AP = A'$ to $AP = PA'$. Bearing in mind that $P = \begin{bmatrix} | & | & | \\ b_1 & b_2 & b_3 \\ | & | & | \end{bmatrix}$ and that $A' = \text{diag}(2, -3, 7)$, this is saying that

$$A \begin{bmatrix} | & | & | \\ b_1 & b_2 & b_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ b_1 & b_2 & b_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{bmatrix} \Longrightarrow \begin{bmatrix} | & | & | \\ Ab_1 & Ab_2 & Ab_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ 2b_1 & -3b_2 & 7b_3 \\ | & | & | \end{bmatrix}$$

This means that $Ab_1 = 2b_1$, $Ab_2 = -3b_2$ and $Ab_3 = 7b_3$, so that $B = \{b_1, b_2, b_3\}$ is a basis of $\mathbb{R}^3$ consisting of *eigenvectors* of A.