Lecture 14: Change of Basis

March 1, 2024

Row rank and column rank

Let A be a $m \times n$ matrix.

The row rank r of A is the dimension of the row space of A, which is the subspace of $\mathbb{R}^n$ spanned by the rows of A.

The column rank c of A is the dimension of the column space of A, which is the subspace of $\mathbb{R}^m$ spanned by the columns of A.

The row rank is at most m and the column rank is at most n, but both can be less.

> Theorem The row rank and column rank are same for every matrix.

So we can just refer to the rank.

Row rank = Column rank

For a $m \times n$ matrix A, write r for its row rank and c for its column rank.

Choose a basis for the row space of A, and write its elements as the rows of a $r \times n$ matrix P.

Since every row of A is a linear combination of the rows of P, there is a $m \times r$ matrix Q for which $A = QP$. (Think about this!)

But now every column of A is a linear combination of the r columns of Q, so the dimension of the column space of A is at most r: $c \leq r$.

To see that $r \leq c$ also, start with a basis for the column space of A, and write its c elements as the columns of a $m \times c$ matrix P'. Then $A = P'Q'$ for some $c \times n$ matrix Q', and every row of A is a linear combination of the c rows of Q'. Hence $r \leq c$.

Since $c \leq r$ and $r \leq c$, we conclude $r = c$.

Coordinates

Lemma If $\{b_1, \dots, b_n\}$ is a basis of a vector space V, then every element of V has a unique expression as a linear combination of $b_1, \dots, b_m$.

Proof Suppose for some $v \in V$ that

$$\begin{aligned} v &= a_1 b_1 + a_2 b_2 + \dots a_n b_n, \text{ and} \\ v &= a_1' b_1 + a_2' b_2 + \dots a_n' b_n, \end{aligned}$$

for scalars a_i, a_i'. Then $0_V = (a_1 - a_1')b_1 + (a_2 a_2')b_2 + \cdots + (a_n - a_n')b_n$. Since B is linearly independent, the coefficients $a_i - a_i'$ are all zero and the two expressions for v are identical. $\square$

Example In $\mathbb{R}^2$, the standard coordinates of $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ are $(4, 3)$. With respect to the basis $\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\}$, the coordinates of $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ are $(2, 1)$. This is saying that

$$\boxed{\begin{bmatrix} 4 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} 2 \\ -1 \end{bmatrix}.}$$

Coordinates with respect to different bases

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

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

Question Given an element of $\mathbb{R}^3$, for example $v = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$, how do we find the B-coordinates of v?

Answer We know $v = 2e_1 - 3e_2 + 4e_3$. We need the B-coordinates of the standard basis vectors e_1, e_2, e_3. Then $[v]_B = 2[e_1]_B - 3[e_2]_B + 4[e_3]_B$.[1]
To find $[e_1]_B$:

$$e_1 = xb_1 + yb_2 + zb_3 \Longrightarrow \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} x + \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} y + \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} z = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & 0 \\ 4 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

This is saying that $[e_1]_B$ is the first column of the inverse of the matrix that has the elements of the basis B as its three columns. Columns 2 and 3 of this inverse are $[e_2]_B$ and $[e_3]_B$.

[1] We write $[v]_B$ for the column vector with the B-coordinates of v as entries

Change of Basis

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

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

How to find the B-coordinates of $v \in \mathbb{R}^3$, for example $v = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$:

Write a 3×3 matrix P whose columns are the basis elements of B.

$$P = \begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & 0 \\ 4 & 0 & 2 \end{bmatrix},\quad P^{-1} = \begin{bmatrix} -\frac{1}{7} & -\frac{2}{7} & \frac{2}{7} \\ 0 & -1 & 0 \\ \frac{2}{7} & \frac{4}{7} & -\frac{1}{14} \end{bmatrix}$$

To find the B-coordinates of any $v \in \mathbb{R}^3$, multiply v on the left by P^{-1}.

$$\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}_B = \begin{bmatrix} -\frac{1}{7} & -\frac{2}{7} & \frac{2}{7} \\ 0 & -1 & 0 \\ \frac{2}{7} & \frac{4}{7} & -\frac{1}{14} \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{12}{7} \\ 3 \\ -\frac{10}{7} \end{bmatrix}$$

This is saying $v = \frac{12}{7} b_1 + 3b_2 - \frac{10}{7} b_3$, which can be checked.

P^{-1} is called the change of basis matrix from the standard basis to B.

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.