Chapter 3: Linear transformations, eigenvectors, similarity

Themes for Chapter 3

- It is useful to be able to move between different bases for a given vector space;
- One basis may be far better than another for describing a particular linear transformation - the standard basis is not always the most useful one;
- Everything can be interpreted in terms of matrix algebra, although the setup takes some work.

Moving between two bases

Suppose we have another basis $\mathcal{B} = \{b_1, b_2, b_3\}$ of R^3 (besides the standard basis $\{e_1, e_2, e_3\}$), where

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

Question: Suppose we have some vector in $\mathbb{R}^3$, for example $v = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

What are the coordinates of v with respect to $\mathcal{B}$?

Another Question: Why would we want to know this?

How to write the $\mathcal{B}$-coordinates of v?

If we knew how to write e_1, e_2 and e_3 as a linear combination of b_1, b_2, b_3, we could do the same for v (or any vector).
The $\mathcal{B}$-coordinates of e_1 are the values of x, y, z in the unique solution of

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \text{or } B \begin{bmatrix} x \\ y \\ z \end{bmatrix} = e_1.$$

The corresponding values are given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = B^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

which means they are the entries of Column 1 of B^{-1}. In the same way, the $\mathcal{B}$-coordinates of e_1 and e_3 are given by Columns 2 and 3 of B^{-1}.

The matrices B and B^{-1}

For our example:

$$B = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ -1 & 2 & 0 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \end{bmatrix}$$

Looking at (for example) Column 2 of B^{-1} we can confirm that its entries are the $\mathcal{B}$-coordinates of e_2.

The change of basis matrix

Now for the $\mathcal{B}$-coordinates of $v = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$. We write $[v]_{\mathcal{B}}$ for the column whose entries are the $\mathcal{B}$-coordinates of v. We can now achieve this through a matrix-vector product.

$$\begin{aligned} v = 2e_1 + 1e_2 + 3e_3 \Longrightarrow [v]_{\mathcal{B}} &= 2[e_1]_{\mathcal{B}} + 1[e_2]_{\mathcal{B}} + 3[e_3]_{\mathcal{B}} \\ &= \begin{bmatrix} [e_1]_{\mathcal{B}} & [e_2]_{\mathcal{B}} & [e_3]_{\mathcal{B}} \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 9/2 \\ 1/2 \end{bmatrix} \end{aligned}$$

Conclusion: $v = 6b_1 + \frac{9}{2}b_2 + \frac{1}{2}b_3$.

Exercise: Confirm this conclusion by direct calculation.

General conclusion

To find the $\mathcal{B}$-coordinates of *any* vector v in $\mathbb{R}^3$, what we need to do is multiply v on the left by the *change of basis* matrix from the standard basis to $\mathcal{B}$. This is the inverse of the matrix whose columns are the elements of $\mathcal{B}$ (written in the standard basis).

Learning outcomes for Section 3.1

1. How to recognize when a set of n column vectors in $\mathbb{R}^n$ (or $\mathbb{F}^n$) forms a basis.
2. To recognize that elements of $\mathbb{R}^n$ (or $\mathbb{F}^n$) have different coordinates with respect to different bases.
3. To use the change of basis matrix to write the coordinates of any vector in $\mathbb{F}^n$ with respect to a given basis.

The Matrix of a Linear Transformation

Let V and W be $\mathbb{F}$-vector spaces and let $\phi : V \to W$ be a linear transformation. Recall what this means:

- $\phi(u+v) = \phi(u) + \phi(v)$ for all $u, v \in V$, and $\phi(\lambda v) = \lambda\phi(v)$, for all $v \in V$ and $\lambda \in \mathbb{F}$.

Example If A is matrix in $M_{m\times n}(\mathbb{F})$, then left multiplication by A defines a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$. $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix}$ defines a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ via

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \to \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2a + 3b + c \\ a - 2b + c \end{bmatrix}.$$

Note that the images of the three standard basis vectors of $\mathbb{R}^3$ under this transfomation are respectively the columns of A.

Every linear transformation is left matrix multiplication

Now suppose that $\dim V = n$ and $\dim W = m$. Once we choose bases $\mathcal{B}_V$ and $\mathcal{B}_W$ for V and W, *every* linear transformation from V to W looks like the one in the last slide.

Example The differential operator D, which sends every polynomial to its derivative, is a linear transformation from $\mathbb{R}[x]$ to $\mathbb{R}[x]$, and from P_3 to P_2.

$\mathcal{B}_3 = \{x^3, x^2, x, 1\}$ and $\mathcal{B}_2 = \{x^2, x, 1\}$ - bases for P_3 and P_2 respectively. Look at its image in P_2 under D of elements of $\mathcal{B}_3$, as vectors with $\mathcal{B}_2$-coordinates.

$$x^3 \to 3x^2 : \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2}, \; x^2 \to 2x : \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}_{\mathcal{B}_2}, \; x \to 1 : \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{B}_2}, \; 1 \to 0 : \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2}.$$

Example: The differential operator $P_3 \to P_2$

The $\mathcal{B}_3$-coordinates of the element $p(x) = ax^3 + bx^2 + cx + d$ are given by the column $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$, and the $\mathcal{B}_2$ coordinates of the derviative of p are given by

$$a\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2} + b\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}_{\mathcal{B}_2} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{B}_2} + d\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}_{\mathcal{B}_3}$$

The 3×4 matrix above is the matrix of D with respect to the bases $\mathcal{B}_3$ and $\mathcal{B}_2$. Its columns are the images under D of the elements of $\mathcal{B}_2$, written with respect to $\mathcal{B}_3$. To apply the operator to any polynomial $p(x)$, we can write it as a column vector (with respect to $\mathcal{B}_3$) and then multiply by the matrix. The result has the $\mathcal{B}_2$-coordinates of $p'(x)$.

Same story, different basis

This matrix depends on the choice of bases! Suppose we keep the basis $\mathcal{B}_2$ of P_2, but take $\mathcal{C}_3 = \{x^3 + x^2, x^2 + x, x + 1, 1\}$ as our basis of P_3. The matrix of the differential operator with respect to this choice has the $\mathcal{B}_2$-coordinates of the derivatives of elements of $\mathcal{C}_3$ as its columns, it is given by

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

To use this matrix to determine the derivative of (for example) $f(x) = x^3 + 4x^2 - x - 2$, first write $f(x)$ with respect to $\mathcal{C}_3$: $1(x^3 + x^2) + 3(x^2 + x) - 4(x + 1) + 2(1)$. Then

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -4 \\ 2 \end{bmatrix}_{\mathcal{C}_3} = \begin{bmatrix} 3 \\ 8 \\ -1 \end{bmatrix}_{\mathcal{B}_2}.$$

Lecture 15: The kernel and Image

$\phi : V \to W$ is a linear transformation of finite dimensional vector spaces.

Definition The kernel of ϕ , denoted $\ker \phi$, is the set of elements of V whose image is the zero vector of W.

$$\ker \Phi = \{v \in V : \phi(v) = 0_W\} \subseteq V.$$

Definition The *image* of ϕ, denoted $\operatorname{image} \phi$, is the subset of W consisting of the images of all the elements of V.

$$\operatorname{image} \phi = \{\phi(v) : v \in V\} \subseteq W.$$

The kernel and image of ϕ are subspaces of V and W.

The nullspace of a matrix

Example For the linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ defined as left multiplication by the matrix $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix}$, the *kernel* consists of all vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ for which

$$\begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In the matrix context, this is referred to as the (right) *nullspace* of A. We can find it by row reduction; in this example it consists of all vectors of the form $t \begin{bmatrix} -5 \\ 1 \\ 7 \end{bmatrix}$ where $t \in \mathbb{R}$ - a subspace of dimension 1 of $\mathbb{R}^3$.

The nullspace of $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix}$ is the linear span of $\begin{bmatrix} -5 \\ 1 \\ 7 \end{bmatrix}$ in $\mathbb{R}^3$.

The image of the "left multiplication by A" linear transformation is the linear span of the three columns of A. In the matrix context, it is called the column space of A. In this example, it is all of $\mathbb{R}^2$, since the first two columns of A (for example) span $\mathbb{R}^2$.

We can note that in this example, the kernel (nullspace) and image (column space) have dimension 1 and 2, and 1+2=3, and 3 is the dimension of the domain $\mathbb{R}^3$. This is not a coincidence, but a case of the Rank-Nullity Theorem.

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*.

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

$$\dim(\ker \phi) + \mathrm{rank}\phi = \dim V.$$

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.

Learning Outcomes for Section 3.2

1. *To recall the definition of a linear transformation*
   as a function between vector spaces that respects the addition and scalar multiplication operations.
2. *To note that left multiplication by any $m \times n$ matrix is a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$,*
   and that the columns of the matrix are the images of the standard basis vectors of $\mathbb{F}^n$
3. *That every linear transformation can be represented as left multiplication by a matrix,*
   For relatively small and manageable examples, you should be able to write down the matrix that does this, and realize that it depends on the choice of basis (we will come back to this point).
4. *To recognize the terms kernel, image, nullspace, nullity, rank and column space.*
5. *To be able to state and interpret the Rank-Nullity Theorem, in its versions for matrices and for linear transformations*